ACCURACY, VALIDITY AND RELEVANCE OF PROBABILISTIC BUILDING ENERGY MODELS

A Dissertation
Presented to
The Academic Faculty

by

Qinpeng Wang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosphy in the
School of Architecture

Georgia Institute of Technology
August 2016

ACCURACY, VALIDITY AND RELEVANCE OF PROBABILISTIC BUILDING ENERGY MODELS

Approved by:

Prof. Godfried Augenbroe, Advisor
School of Architecture
Georgia Institute of Technology

Dr. Pieter De Wilde
School of Architecture
University of Plymouth, UK

Dr. C. F. Jeff Wu
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Chris Paredis
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Jason Brown
School of Architecture
Georgia Institute of Technology

Date Approved: June 22, 2016

To my parents who have always supported me

ACKNOWLEDGEMENTS

I'm deeply indebted and grateful to my advisor Godfried Augenbroe, for never turning down my meeting request. His sharp insight as well as witty humor, unwavering encouragement as well as constructive criticism nurtured me throughout my study. I would like to extend my deepest gratitude to my committee members: Professor C.F. Jeff Wu, Chris Paredis, Pieter De Wilde, and Jason Brown, for their support and enlightening discussions towards this work.

Special thanks to Yuming Sun, Qi Li, Li Gu, Bass Abushakra, Mike wolf, Michael Street, Michael Wetter, Thierry Nouidui, Bin Yan, Huafen Hu, Shan Ba, Alexander Jordan, Professor Tilmann Gneiting, and Professor Brani Vidakovic, who contributed in discussions that eventually shaped this dissertation. Many thanks to Professor Russell Gentry, Professor Baabak Ashuri, Professor Craig Zimring, Professor John Peponis, Professor Michael Gamble, and Professor Charles Eastman, who once shared their wisdom and knowledge with me in class.

I'd also like to extend gratitude to my lovely colleagues for their company in the lab: Yuna, Roya, Atefe, Yifu, Jihyun, Jaeho, Haoran, Sanghoon, Fei, Zhengwei, Yiyuan, Tianyao, Te, Di, Yijia, Pei, Jige, and Jianli. A massive shout out to my friends: Chen, Dan, Zhichen, Shun, Zhuzhou, Yining, Huizhi, Cong, Junying, and all others, because they are swag.

Last but not least, I would not have accomplished all this without the full support and selfless love of my parents, and aunt Jing in Atlanta: I'm sincerely grateful.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
SUMMARY	xiii
CHAPTER	
1 INTRODUCTION	1
1.1 The Need for Uncertainty Analysis	2
1.2 Literature Review	4
1.3 Dissertation Structure	9
2 METHODOLOGY FOR PERFORMING UA	12
2.1 Common Uncertainty Analysis Framework	12
2.2 Uncertainty Quantification Methodology	14
2.2.1 Parameter Uncertainty Quantification	14
2.2.2 Model Form Uncertainty Quantification	15
2.2.3 Uncertainty Propagation	16
2.2.4 Risk Measures and Risk Preference	20
2.2.5 Sensitivity Analysis	20
2.2.6 Iterative Uncertainty Analysis	21
3 UNCERTAINTY TAXONOMY AND REPOSITORY	23
3.1 Uncertainty Taxonomy	23
3.2 Uncertainty Quantification Repository	30
3.2.1 Meso-Scale	30

3.2.2 Building Level 31
3.2.3 HVAC System Level 32
3.2.4 Occupant Level 32
3.2.5 Georgia Tech Uncertainty and Risk Analysis Workbench 32
4 GAP ANALYSIS AND QUANTIFICATION 35
4.1 Infiltration 35
4.1.1 Air Leakage Area 37
4.1.2 Air Leakage Distribution 43
4.2 Workmanship Issues at a Thermal Bridge 45
4.2.1 Methodology 46
4.2.2 Results 49
4.3 Occupancy Variables 51
4.3.1 Proposed Modeling Framework 54
4.3.2 Results and Discussions 61
4.3.3 Relevance of Stochastic Occupancy Models 67
4.4 HVAC System Uncertainty 68
4.4.1 Literature Review 71
4.4.2 Schematic of an Air-Handling Unit with Terminal VAV Boxes 73
4.4.3 Description of the EnergyPlus Model and Potential Inadequacy 75
4.4.4 Description of the High Fidelity Model 77
4.4.5 Uncertainty Quantification of the EnergyPlus VAV Model 80
4.4.6 Result Generalization for Future Analysis of VAV Systems 92
4.4.7 Realization Uncertainty of Duct System Resistance 94
4.4.8 Conclusions 101
5 SENSITIVITY ANALYSIS 103

5.1 Recap on Variance-Based Methods 104
5.1.1 Features of First Order Effects 105
5.1.2 Connection with Standardized Regression Coefficients 105
5.2 Group Lasso with Discrete Categorical Variables and SLHS 106
5.2.1 Consideration of Categorical Variables 106
5.2.2 Grouping of Variables with Group Lasso 107
5.2.3 Sliced Latin Hypercube Sampling 108
5.3 Case Study 108
5.3.1 Sensitivity Analysis Description 109
5.3.2 Results and Discussions 110
5.3.3 Iterative Uncertainty Analysis 111
6 VERIFICATION OF PROBABILISTIC ENERGY PREDICTIONS 113
6.1 Conventional Approach 113
6.2 Methodology 116
6.2.1 Formal Definitions 117
6.2.2 Interpretations 118
6.2.3 Diagnosis for Probabilistic Calibration 120
6.2.4 Ranking of Competing Forecasts 120
6.2.5 Simulation Examples 121
6.2.6 HVAC Uncertainty Factor 123
7 VALIDITY OF PROBABILISTIC ENRERGY PREDICTIONS 136
7.1 Contractual Decisions in a Performance Contract 136
7.1.1 Axiomatic Utility Theory 138
7.1.2 Impact of Risk Information on Decisions 138
8 EMPIRICAL SPECULATION ABOUT THE RELEVANCE OF UA 151

8.1 Quantification of Impact 152

9 CONCLUSIONS AND FUTURE WORK 156

9.1 Summary and Conclusions 156

9.1 Recommendations for Future Study 157

APPENDIX A 160

REFERENCES 166

LIST OF TABLES

	Page
Table 2.1: Example of performance indicators and risk measures	21
Table 3.1: Material property uncertainty	32
Table 4.1: Flow exponent from the literature	40
Table 4.2: R and X in literature	44
Table 4.3: Material property and dimensions used in the hypothetical case	48
Table 4.4: Detailed temperature distribution along the internal wall surface	50
Table 4.5: Inferred temperature distributions	51
Table 4.6: Statistics of regression models	64
Table 4.7: Polynomial coefficients for fan performance curves	77
Table 4.8: Pressure loss of the critical paths for both actual and design situations	102
Table 5.1: Categorical variables considered in the case study	109
Table 6.1: Building general information summary	125
Table 7.1: Cost breakdown by energy conservation measure	141
Table 7.2: Results of ECM packages	148
Table A.1: Total pressure loss calculations by sections for the actual situation	160
Table A.2: Loss coefficient summary by sections for the actual situation	162
Table A.3: Total pressure loss calculations by sections for the design situation	163
Table A.4: Loss coefficient summary by sections for the design situation	165

LIST OF FIGURES

Page

Figure 1.1: Simulation as a virtual experiment 2

Figure 2.1: Common uncertainty analysis framework 14

Figure 2.2: Typical procedures for a Monte Carlo analysis 17

Figure 2.3: An example LHS 19

Figure 3.1: Physical model uncertainty EER diagram 24

Figure 3.2: Scenario EER diagram 27

Figure 3.3: Different scales of uncertainty 30

Figure 3.4: Workflow of GURA-W 33

Figure 4.1: Leakage flow rates (m3/s.m^2) at reference pressure difference of 75 Pa 38

Figure 4.2: Kernel density estimate of the pressure exponent n 39

Figure 4.3: A hierarchical model 41

Figure 4.4: Posterior distribution of predicted $log(ELA)$ 42

Figure 4.5: Posterior distribution for population mean air leakage (m^3/h.m^2) at 75 Pa 43

Figure 4.6: A construction detail with no thermal bridging effect 47

Figure 4.7: "Proper workmanship" (upper) and "Bad workmanship" (lower) scenarios 49

Figure 4.8: Occupancy information available to the modeler (from low to high) 53

Figure 4.9: A basic hierarchical model 57

Figure 4.10: Propose framework for a meta-analysis 58

Figure 4.11: Occupancy sensor connected with a minicomputer above an entrance 62

Figure 4.12: Presence (left) and lighting/appliance (right) profile ground truth 63

Figure 4.13: De-trended presence (left) and lighting/appliance (right) variability 64

Figure 4.14: Diagnosis plots for the $AR(1)$ model for X_t 64

Figure 4.15: Cross-correlation between Y_t' and $\widehat{a_t}$ 65

Figure 4.16: Mean occupancy profiles: presence (left) and lighting/appliance (right) 66

Figure 4.17: A typical presence (left) and lighting/appliance (right) profile 66

Figure 4.18: Schematic for control sequence VAV 2A2-21232 74

Figure 4.19: Fan characteristics in this study 78

Figure 4.20: Configuration of EnergyPlus with schedules 81

Figure 4.21: A specialized component in Modelica 83

Figure 4.22: Overall Modelica model for the VAV system 84

Figure 4.23: Western VAV box operation 85

Figure 4.24: Drawback of the current control sequence 86

Figure 4.25: Comparison of core zone room air temperature 87

Figure 4.26: Comparison of supply air volume flowrate 88

Figure 4.27: Detailed temperature and supply air flowrate profile 89

Figure 4.28: Comparison of fan power 89

Figure 4.29: Curve fitting results 90

Figure 4.30: Comparison of fan power with updated coefficients 91

Figure 4.31: Predicted fan operating points by both models 92

Figure 4.32: Comparison of fan power for an updated building in Phoenix 93

Figure 4.33: Comparison of fan power for updated system resistance in Atlanta 94

Figure 4.34: Predicted versus actual system pressure loss 95

Figure 4.35: Supply air ductwork (highlighted) on the first floor 98

Figure 4.36: Design (upper) and actual (lower) duct layout 99

Figure 4.37: Additional elbow at an end branch and compressed ducts 101

Figure 5.1: Sensitivity index for heating (upper) and cooling (lower) in Atlanta 111

Figure 5.2: Results of original and refined uncertainty analysis 112

Figure 6.1: True distribution versus measurements (left) and NBI data (right) 116

Figure 6.2: CRPS for a single observation 121

Figure 6.3: Realizations against prediction by both forecasts 122

Figure 6.4: Probabilistic (upper) and marginal (lower) calibration results by forecasts 123

Figure 6.5: Georgia Tech alumni association 127

Figure 6.6: Georgia center for advanced telecommunications technology (GCATT) 128

Figure 6.7: Office of human resources 128

Figure 6.8: Ivan Allen college of liberal arts 129

Figure 6.9: Language institute O'Keefe building 130

Figure 6.10: Institute of paper science and technology (IPST) 130

Figure 6.11: Preliminary calibration results: probabilistic (left) and marginal (right) 132

Figure 6.12: Calibration results without IPST: probabilistic (left) and marginal (right) 133

Figure 6.13: Calibration results with HVAC factor: probabilistic (left), marginal (right) 134

Figure 7.1: Assumed NPV distribution by a trusting and a most mistrustful client 145

Figure 7.2: Expected utility for ESCOMP (left) and the school board (right) 147

Figure 7.3: ESCOMP’s expected utility with proposed guaranteed savings 149

Figure 8.1: Consumer risk premium distribution: baseline (left), high demand (right) 153

Figure 8.2: Risk premium with risk information: baseline (left), high demand (right) 153

Figure 8.3: Commercial delivered energy intensity projections 155

SUMMARY

Residential and commercial buildings consume 41% of total U.S. energy consumption. Since improving energy efficiency is still the most cost efficient energy saving option in the U.S., it is not surprising that many new buildings represent a push towards ultra-efficiency. Many studies argue that this calls for the use of high fidelity prediction models that by necessity will be probabilistic in nature due to many sources of uncertainty that affect the translation of a design specification into the actual reality of a constructed and operated facility. To inspect the fidelity of these probabilistic models against traditional deterministic models, we pose questions that address three major aspects of this new generation of building energy models:

- Accuracy: do these models give more "correct" answers?
- Validity: do these models lead to "better" design/retrofit decisions?
- Relevance: does a profession that deploys these models provide "higher" value to the industry?

This dissertation addresses the first question by identifying gaps in our understanding and quantifying various sources of model uncertainty reported in recent literature. Insufficiently understood and not well-quantified sources are further studied and resolved. The results of the above are analyzed in a sensitivity analysis that ranks input parameters alongside with model form uncertainties. Next, we adapt proven methods to conduct verification of probabilistic building energy models. Probabilistic calibration, marginal calibration and a continuous rank probability score are used to evaluate the "*correctness*" of the new generation of models. We illustrate the challenges of delivering *validity* proofs in a case study where outcomes of uncertainty analysis are translated into (monetary) risks and their influence is analyzed in a decision-making scenario involving energy performance contracts. Lastly, the study introduces a

speculative approach to proving *relevance* by quantifying the overall societal benefit of a transparent risk framework that has the potential to unlock currently stagnating capital flow into large-scale building retrofits.

CHAPTER 1

INTRODUCTION

In 2014, residential and commercial buildings consumed about 41% of total U.S. energy consumption, or about 40 quadrillion British thermal units (EIA, 2015). Since energy efficiency is holding steady as the least-cost energy resource option (measured by levelized cost of electricity) in the U.S. (Molina, 2014), more and more attention has been drawn to new designs that incorporate the consideration of energy efficiency in the design process. This trend is demonstrated by enhanced prescriptive specifications in building codes and regulations such as ASHRAE standard 90.1, and the fact that 675.9 million square feet of real estate space became LEED certified in 2014 with a growth rate of 13.2% from 2013 (U.S. Green Building Council, 2015). It is noteworthy that buildings are complex artifacts, and as new designs push the envelope of building performance, their performance evaluation has to be backed up by building energy simulation.

The role of building energy simulation in the design and engineering of building has been established since 1960s. Since then, many tools have matured and proliferated into the consultants' offices worldwide. Building energy simulation deals with the energy and mass flow in the built environment. The "behavioral model" of a building is created, given the stage of its development. The whole spectrum of the development stage could involve as-designed, as-built and as-operated specifications. Throughout these three stages, the availability and certainty of information relevant to the building deepens. This dissertation mainly looks at predictions at the design stage, but the discussion partially involves the latter two stages as well. By providing insight towards the consequences of design decisions, building energy simulation informs the engineering design process.

Taking one step back, we realize that models are idealizations of the real physical world, with some level of abstraction. This idealization is represented by mathematical

formulas governed by physical conservation laws. Then think of simulation as a virtual experiment (Figure 1.1), which subjects this idealization to all kinds of boundary conditions and scenarios of usage, such that some useful conclusions can be drawn from these experiments. Since 1970s, techniques behind our tools gradually matured. Typically, finite difference and finite element methods are adopted as an approximation to partial differential equations (Augenbroe, 1986). The resulting system is a set of differential algebraic equations (DAE) derived through space-averaged treatment of the laws of thermodynamics.

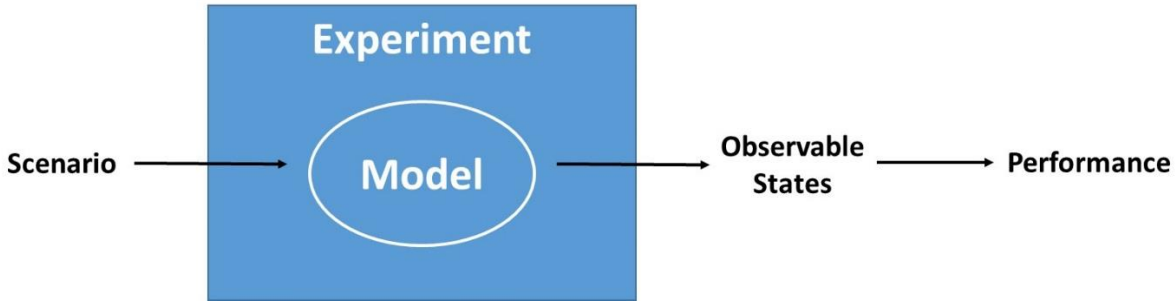


Figure 1.1 Simulation as a virtual experiment

1.1 The Need for Uncertainty Analysis

Despite the development of current energy simulation practice, many challenges remain before the discipline reaches the level of maturity that its growing role in influencing design decisions demands. One of the challenges is closing the so-called "performance gap", or minimizing the discrepancy between prediction and measurement. As such, the assumption that our models are adequate for guiding new designs and planning retrofits cannot be taken for granted. Recent studies (de Wilde, 2014; Ryan & Sanquist, 2012; Turner, Frankel, & Council, 2008) emphasize this point. As we realize that our confidence associated with a one-point or deterministic prediction of energy is low, two trains of thought emerge on how to deal with this situation. The traditional approach is to compare the prediction with the measurement of many realizations of the same product, which leads to some measure of correctness. However, we only realize

every building one time, so it will be very hard to generalize the single measurements we have for every realization. Even worse, if prediction and measurement are far apart, it is probably the result of many compounding effects that we cannot separate and learn from. The other approach relies on uncertainty analysis (UA) as a vehicle for analyzing how wrong our predictions are. An additional benefit of uncertainty analysis is that, as we study uncertainty, we raise the understanding of the physical world and gain better insights in how to manage the most influential sources of uncertainty.

Another need for uncertainty analysis is in the quantification of risk. As mentioned before, uncertainty analysis estimates the level of confidence that can be placed in our predictions. Without such a measure of "correctness", it is impossible to judge the fitness of the prediction for making decisions. In other words, trusting a deterministic model output associated with high uncertainty may lead to considerable risk to the decision maker. Risk is a measure of the probability and consequence of uncertain future events. It is the chance of an undesirable outcome. That outcome could be a loss or a potential gain that is not realized (such as an investment that does not produce expected benefits). Given the presence of a hazard or an opportunity, there are two important components to a risk: probability of an event and the undesirable outcome or consequence linked to that event. Risk is often described by the simple equation:

$$\text{Risk} = \text{Probability} \times \text{Consequence}$$

Consider this expression as a mental model that helps us think about risk rather than an equation that defines it. What this expression is conveying is not so much that this is the manner in which risks are calculated (they are not) as much as that both of these elements must be present for there to be a real risk. If an event of any consequence has no probability of occurrence, there is no risk. Likewise, if there is no consequence or undesirable outcome, there is no risk either.

Consequently, this dissertation takes the uncertainty analysis approach. The deterministic prediction is replaced with a probabilistic prediction, and therefore any

downstream processes such as verification and decision making have to work with probabilistic distributions. In an elementary uncertainty analysis, we start from some crude analysis of the uncertainties that reside in some model parameters, and then we propagate all of these identified uncertainties, through a Monte Carlo engine that runs over and over with different sampled values of these uncertainty parameters. By recombining the outcome of all these sample runs, we construct a distribution of the variable of interest. The elementary uncertainty analysis, of course, has plenty of room for improvement, but before that, we move on to a survey of the historical research in the area of uncertainty analysis.

1.2 Literature Review

Previous attempts at measuring the correctness of models compare the realizations of design with deterministic computational model outputs. Such an approach is impossible to generalize and it is hard to separate many confounding factors that contribute to the discrepancy. Early work by, for instance, Macdonald (2002) and de Wit (2001) recognize the role of uncertainty analysis in quantifying the degree of inaccuracy associated with our model answers. Uncertainties with model input parameters are identified and then propagated with proper sampling design such as Latin Hypercube sampling, into the outcome for an uncertainty range of the variable of interest. Specifically, Macdonald and Strachan (2001) provide an insightful review about sources of uncertainty in thermal predictions and the implementation of uncertainty analysis in the program ESP-r, whereas de Wit and Augenbroe (2002) combine a crude assessment of plausible ranges for model parameters and an expert judgement study to evaluate uncertainty in wind pressure coefficients in a more rigorous way, followed by analysis of their potential impact on design decisions with an explicit recognition of the preference of decision makers. Following their pioneering work, lots of research in this space has provided evidence that uncertainty analysis needs to be an intrinsic part of quantifying

how wrong our model outcomes can be. Burhenne, Tsvetkova, Jacob, Henze, and Wagner (2013) propose a framework to enhance decision-making at the design stage by quantifying uncertainty associated with both building simulation and cost-benefit calculations, and Monte Carlo filtering is applied to determine the design space that leads to a positive net present value of investment. Corrado and Mechri (2009) conduct uncertainty and sensitivity analysis for building energy rating on a dwelling with the quasi-steady simplified monthly method in the ISO 13790:2008 standard, and concludes that the asset energy rating is only slightly affected by input parameter uncertainties. Chalabi et al. (2015) subject the physical model of a dwelling to uncertainty analysis and studies health impacts of housing policies and housing energy interventions in UK. Daly, Cooper, and Ma (2014) identify sensitive simulation inputs for total predicted building energy consumption in two Australian office buildings and finds the predicted energy consumption for the archetypes varies by more than 50% using high and low assumptions for all Australian capital cities. Eisenhower, O'Neill, Fonoberov, and Mezić (2012) present a method that extends the capability of traditional sensitivity analysis that focuses on about 1,000 parameters. Gang, Wang, Shan, and Gao (2015) take into account cooling load calculation uncertainties and explore a cooling system sizing and configuration method that can determine the appropriate system capacity with quantified confidence. Gaterell and McEvoy (2005) retrospectively examine the performance of insulation measures in an existing residential dwelling and study the impact of projected climate change uncertainties, and conclude that their effect could be considerable. Hopfe, Augenbroe, and Hensen (2013) discuss how design decision making can be based on uncertainty assessments and apply analytic hierarchy process including uncertainty information for a rational choice between two HVAC system designs with regard to energy efficiency and thermal comfort. In a separate paper, Hopfe and Hensen (2011) explore the benefits of the integration of uncertainty analysis in building performance simulation, namely, model simplification, robustness assessment, quality assurance, and

decision support. Huang, Huang, and Wang (2015) propose a prototype of HVAC system design under uncertainty using multiple-criterion decision-making technique, which can assess system performance at the design stage in terms of multiple performance indicators and the customers' requirements and preferences. Kim and Augenbroe (2013) argue that performance evaluation of supervisory demand-side control cannot go without explicit recognition of uncertainty associated with the prediction of the power demand. Liao, Huang, Sun, and Zhang (2014) identify and analyze the impacts of the uncertainties associated with building instantaneous cooling load prediction on several chiller sequencing control strategies. Maasoumy, Razmara, Shahbakhti, and Vincentelli (2014) recognize model uncertainty as a significant challenge to the application of model predictive controllers (MPC) and proposes methodologies for parameter estimation and controller type selection. Menassa (2011) quantifies the value of energy efficiency investment for existing buildings by taking into account uncertainties associated with the life cycle cost-benefit analysis of the investment and augmenting traditional valuation method with modern option pricing theory. Dominguez-Munoz, Cejudo-Lopez, and Carrillo-Andres (2010) propagate uncertainty with the input data through the building model and determine the impact on the peak cooling load, which is of significant consequence to HVAC engineers during system design. Parys, Breesch, Hens, and Saelens (2012) investigate the feasibility of passive cooling in new office buildings in Belgium under uncertainty in terms of two passive cooling schemes, and recommend heat gain limiting targets and measures. Silva and Ghisi (2014a) argue that despite physical, occupation, and weather uncertainties, modeling uncertainty (modeler's bias in our terminology) should play a role by comparing models of the same building that differ in relation to external geometry, groups of internal zones and internal thermal mass. Additionally, Silva and Ghisi (2014b) analyze uncertainties and conduct sensitivity analysis in the performance evaluation of a low-income house in Brazil with an emphasis on user behavior and physical parameters. Smith, Luck, and Mago (2010) study a

combined cooling, heating and power system and compare with conventional systems under various operating strategies with input and model data uncertainty with a case study using a small office building in Atlanta. Spitz, Mora, Wurtz, and Jay (2012) provide a practical application of uncertainty and sensitivity analysis on an experimental house in France, successively applying local sensitivity analysis, correlation analysis, uncertainty analysis and global sensitivity analysis. Srivastav, Tewari, and Dong (2013) propose the modeling of building baseline performance based on Gaussian Mixture regression (GMR) with parameterized and locally adaptive uncertainty quantification. Tian and de Wilde (2011) identify the key to the assessment of the adaptability and resilience of building to changing climate conditions as the uncertainty and sensitivity analysis of building thermal performance in the long term future (50 to 80 years) with a case study using the probabilistic climate change projections for the United Kingdom (UKCP09 dataset). Wang, Mathew, and Pang (2012) study uncertainty associated with the performance of a medium-size office building with an emphasis on actual weather and building operational practices, and find that uncertainty range due to operational factors can hinder the effect of efficiency design features. Most of the above work focuses on our incomplete knowledge of model input parameters, without sufficient recognition of the role of model inadequacy. We conclude from the ongoing work and emerging results that statements of performance that reflect uncertainty will play an increasingly important role in the decision making process of future building owners and occupants. In addition, we argue that uncertainty analysis paves the road to higher fidelity, because as we study uncertainty, we raise the understanding of the physical world and capture it with models, such that better insights on where we need to make improvements can be obtained.

As we are determined to push uncertainty analysis to a new level, the high performance building group at Georgia Tech worked on the NSF funded EFRI-SEED project: Risk-Conscious Design and Retrofit of Buildings for Low Energy, which lasted

from 2010 through 2015. This project targets the theories, models and tool base that enhance our capability to predict future energy consumption and enable a transparent assessment of the risks in the use of certain technologies.

The more recent development of uncertainty analysis resulting from the EFRI-SEED project involves several improvements over the approaches cited above. The approach still employs the basic uncertainty analysis workflow, but has become better in quantifying uncertainties that reside in these parameters. Furthermore it explicitly recognizes the many hidden discrepancies in the models that we use, and which cannot be exposed as parameters uncertainty. This has been categorized as model form uncertainty (Sun, 2014), which is similar to what most of the statisticians would call model discrepancy, and it is very important to identify them. The other major category in UA which is present in any prediction, but has received more explicit emphasis recently is scenario uncertainty, where the most important contributors are obviously associated with internal and external conditions over the course of the virtual experiment, most notably occupant, control and weather dynamics. Recent work by Sun, Su, Wu, and Augenbroe (2015) proposes the consideration of model discrepancy in assessing the accuracy of model outcomes, with an emphasis on the model of solar diffuse irradiation on tilted planes. Another basic contribution of the EFRI-SEED project are methods and tools. It established a repository containing generic (also referred to as "vanilla") distributions for uncertain parameters in building energy input models. It also produced a software platform or workbench using Java that automatically implements the process of uncertainty analysis, from sampling, to batch simulation and result processing. With this groundwork done, we are equipped to ask three ambitious and hard questions that goes beyond the original intention of the EFRI-SEED project. These questions are important because without providing a satisfactory answer to them, we cannot assess to what degree our work has added significantly to the knowledge of the field:

- Accuracy: do these models give "correct" answers?

- Validity: do these models lead to “better” design/retrofit decisions?
- Relevance: does a profession that deploys these models provide “higher” societal value?

Traditional efforts mainly focus on improving accuracy and extending functionality of our toolset and assume that the answer to the next two questions is affirmative by default. The consequence is that we have no clear understanding how well we are served by these tools, or whether they need to be more (or less) accurate in order to add to validity and relevance. It is worth pointing out that the validity question cannot be circumvented because the validity of a model explicitly verifies how people make the right (or better) decisions when using the model.

1.3 Dissertation Structure

Answering all three questions above in one dissertation seems ambitious, and especially some of them require further clarification and qualification. We put major effort of this dissertation in the accuracy question, by filling gaps in our quantification and verification of the correctness and completeness of model uncertainty. For the question of validity, interpreting the validity of UA as “leading to better decisions” requires an approach to validity research that is largely unproven or even unexplored. For the question of relevance, we will only show an empirical speculation. As such, we split the dissertation into three parts.

Part 1 analyzes the current uncertainty quantification work and identifies the gaps in our understanding and quantifications of model uncertainty reported in literature. These insufficiently understood and not well-quantified sources will be further studied and resolved, potentially with an iterative methodology. Sensitivity analysis is used to identify areas where more focus is required.

Part 2 adapts proven methods for the verification of models in other domains than buildings. We present a road towards rigorous verification of probabilistic predictions

with the aim to prove the ranges of model uncertainties that produce the predictions. Several techniques can substantially help evaluate the correctness of our probabilistic model predictions. Part 1 and 2 together address the accuracy question.

Part 3 deals with the validity and relevance question. This part reveals the difficulty of validity proofs in a case study by translating outcomes of uncertainty analysis into risks and analyzing their influence in the decision-making scenarios. For validity proof, we would have to run experiments with decision makers that use our approach and see how confident they are in making decisions. Because of time and resource constrain, the treatment is limited to one test case where we believe uncertainty and risk analysis add value to the decision making process. Then for the remaining question of relevance, we don't pretend it can be fully addressed in this dissertation. The reason for this is that the theoretical basis as well as large-scale data for the proof is lacking. However, we do not want to leave the relevance question untouched albeit we will do that briefly, and in an intuitive way based on an empirical speculation as to the relevance of uncertainty analysis on the magnitude of overall social benefit of a transparent risk quantification at large.

In summary, the dissertation will be structured as the following:

Chapter 2 Methodology

Part 1

- Chapter 3 Uncertainty taxonomy and repository
- Chapter 4 Gap analysis and quantification
- Chapter 5 Sensitivity analysis

Part 2

- Chapter 6 Verification of probabilistic energy predictions

Part 3

- Chapter 7 Validity of probabilistic energy predictions
- Chapter 8 Empirical speculation to the relevance of UA

- Chapter 9 Summary and conclusions

CHAPTER 2

METHODOLOGY FOR PERFORMING UA

This chapter will prepare the reader with the methodology and techniques typically used in the practice of uncertainty analysis.

2.1 Common Uncertainty Analysis Framework

Quantitative uncertainty analysis typically involves the following key components and procedures (de Rocquigny, Devictor, & Tarantola, 2008):

- A physical model as the idealization of the building of interest
- A variety of sources of uncertainty that reside in parameters of the model
- Decision making scenarios that motivate the uncertainty analysis

The computational model that mimics the reality can be viewed conceptually as a numerical function that links inputs (either fixed or uncertain) to outputs. Formally, if we denote the outputs as $\boldsymbol{z}$, fixed inputs as d, and uncertain inputs as $\boldsymbol{x}$, the inputs and outputs can be linked by a deterministic function $\boldsymbol{z} = G(\boldsymbol{x}, \boldsymbol{d})$. The uncertain model inputs $\boldsymbol{x}$ require rigorous uncertainty quantification (UQ) of their range or probability distribution. Some model inputs $\boldsymbol{d}$ may be fixed, for instance, they might be fully controlled, known to have negligible or secondary impact on the outcome, or they are fixed intentionally for comparative purposes. The choice of model output $\boldsymbol{z}$ depend on variables that are of particular interest in the decision-making process. However, once we acquire an analytical or empirical distribution for the variables of interest and performance indicator, we may need to process them to help us make decisions, as uncertainty is meaningless useless we derive something with it. The decision-making process typically involves some risk measures on which to express risk tolerance (risk preference), for instance, percentages of variability in the performance indicator,

expected value, confidence intervals and quantiles, and probabilities of exceeding a threshold.

With identified uncertainties with model parameters, we compute variables of interest with uncertainty propagation. Such a step is typically carried out using a Monte Carlo (MC) approach with an appropriate sampling technique (e.g. Latin hypercube sampling) in order to increase computational efficiency. The sensitivity analysis is also a very important ingredient in the common framework of UA. The concept of sensitivity analysis is very broad that may refer more generally to certain elementary treatments, such as one-at-a-time variations of the inputs of a deterministic model or partial derivatives. In the context of this dissertation, we use the term sensitivity analysis to refer to the computation of so-called sensitivity indexes for uncertainty parameters x with respect to a given performance indicator. Typical techniques include screening (Morris), regression-based methods (standardized regression coefficient), variance-based methods and so on. In the context of a variance-based sensitivity analysis, the sensitivity index represents the expected reduction in the percentage of output variance, if the parameter under investigation could be known or fixed.

Depending on the goals motivating the uncertainty analysis, there may be a feedback process after the preliminary study. If risk measures do not meet a certain criterion or the uncertainty associated with the variable of interest is too large for comparative decision making, an iterative process may be justified, for instance to improve measurements over certain parameters that show the largest sensitivity to the variable of interest, shift to another scenario, or manage outcome uncertainties with any other means.

Figure 2.1 summarizes the common uncertainty analysis framework outlined above. We will zoom into some of the important procedures in the remaining sections of this chapter.

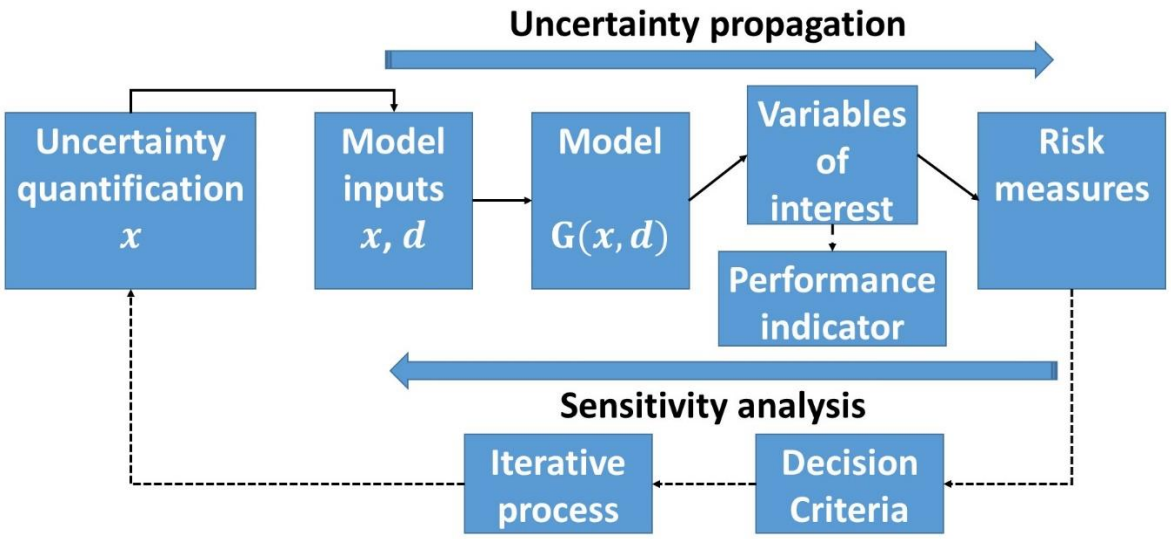


Figure 2.1 Common uncertainty analysis framework (adapted from (de Rocquigny et al., 2008))

2.2 Uncertainty Quantification Methodology

If we consider the modular structures of a building energy model, we realize that input parameters are at the bottom of the pyramids, and cannot be further divided. Therefore, when we deal with the inaccuracy of a module or sub-model as a whole, we face model form uncertainty, whereas when we deal with the lack of knowledge of a specific parameter, we face input parameter uncertainty. Morgan (2009) suggests a differentiation between (1) uncertainty about the value of empirical quantities that appear as parameters in modeling systems, and (2) uncertainty about the model functional form itself.

2.2.1 Parameter Uncertainty Quantification

The uncertainty quantification of input parameters are targeted to represent $\boldsymbol{x}$ with probabilistic distributions. Common probability distributions include uniform distribution, triangular distribution, normal distribution, log-normal distribution, Poisson distribution, exponential distribution, and so on. Regarding probability, the frequentist

(classical interpretation) considers $\boldsymbol{x}$ as observable realizations of uncertain events, and assumes the frequency records of them would allow the inference of probability distributions. The other view involves a quite different interpretation. A subjective interpretation considers the probability distributions as a reflection of the decision-maker subjective preferences following a 'rational preference' set of axioms, or the degrees of belief about these variables without necessarily referencing to frequency observations of them. These competing interpretations of a standard probabilistic setting are quite different, but they imply rather similar practical implementation features. In this dissertation, we do not go overboard with providing a clear distinction between these interpretations, but in general, when we fit probability distributions to collected data around one physical phenomenon within a particular building, the interpretation of probability leans towards the former, while when empirical data are collected from a pool of similar buildings, we interpret it as the basis for our subjective prior belief.

2.2.2 Model Form Uncertainty Quantification

No model is a perfect representation of the physical reality, in other words, model discrepancy is always present. If a sub-model is based on a low fidelity description of the governing physics, we capture this deficiency with model form uncertainty. Let $G(\boldsymbol{\mu})$ denote the output of a module when input parameters take values $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)$. In order to quantify the deficiency of this particular module, we need empirical data from physical experiments or outputs from a higher fidelity module (most likely not implemented in the first place maybe because of complexity or run time inefficiency) as a benchmark. The model discrepancy then is defined as the difference between the benchmark and the module output $G(\boldsymbol{\mu})$, given the same input parameters $\boldsymbol{\mu}$.

$$e = z_{bench} - G(\boldsymbol{\mu})$$

Note that in this section, the benchmark value and module output are both treated as deterministic. The rationale behind this assumption is the following: first, empirical

data collected from a controlled experiment at a submodule scale are typically associated with little uncertainty other than measurement uncertainty (second-order effect compared to e); second, in the quantification of model discrepancy, we assume $\boldsymbol{\mu}$ is known with perfect knowledge, which removes uncertainty from a lack of knowledge of the precise parameter values. As such, the error term e can be modeled with a statistical model, either with only the original input parameters $\boldsymbol{\mu}$, or with the help of additional parameters $\boldsymbol{\nu}$, for a better representation of the physics. Once the error term has been characterized, the sum of the module output $G(\boldsymbol{\mu})$ and $e(\boldsymbol{\mu}, \boldsymbol{\nu})$ will be a close approximation of the benchmark value. For instance, Sun, Heo, et al. (2014) used a high-order meteorological model as the high fidelity model to quantify the uncertainty in a reduced order model of building microclimate. Section 4.4 of this dissertation will also present an example of the quantification of model form uncertainty.

2.2.3 Uncertainty Propagation

Once sources of uncertainty associated with model parameters are quantified, we propagate them through the simulation engine in a Monte Carlo (MC) fashion with an appropriate sampling technique (e.g. Latin hypercube sampling). We will introduce both techniques in this section.

2.2.3.1 Monte Carlo Methods

Monte Carlo methods generally rely on repeated random sampling to obtain numerical approximations to analytical results that are too difficult to solve. The name Monte Carlo actually comes from the gaming tables at the casinos of Monte Carlo, as lots of games are played by chance. Monte Carlo methods are mainly used for optimization, numerical integration and generating draws from a probability distribution.

Monte Carlo methods generally take the steps as presented in Figure 2.2. We begin by initializing the model. We then repeat the procedure many times, in which we

pick a set of random numbers from which we derive values for the input parameters in our model. In each iteration, we collect and store output data from the model such that we can aggregate and analyze them later. For instance, we can construct the entire pdf of the output variable with the histogram of stored model outcomes.

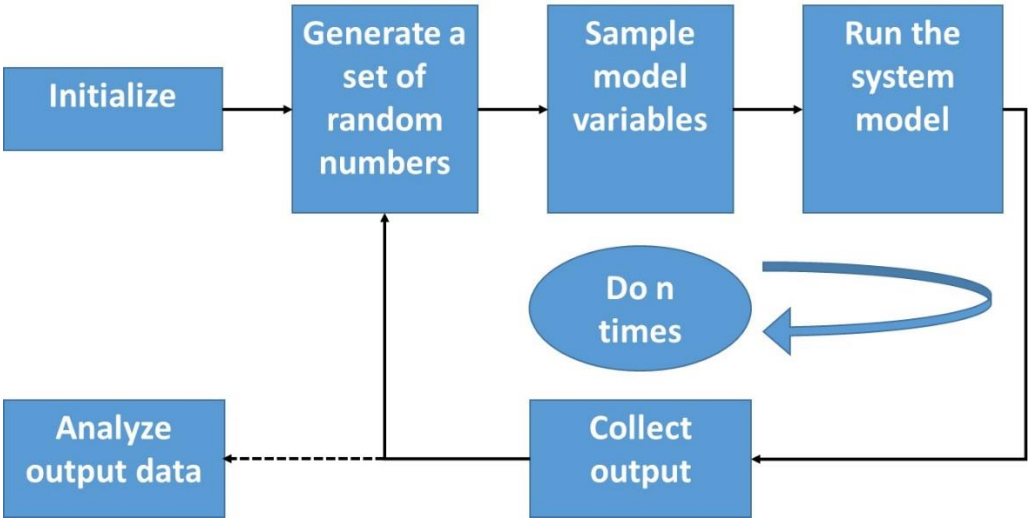


Figure 2.2 Typical procedures for a Monte Carlo analysis

In our application, Monte Carlo integration is used most often, as an expected utility (a well-known concept in decision theory) and a CDF (cumulative density function) value can both be interpreted as integrals. For an example of Monte Carlo integration, we suppose the objective is to compute $E\big(f(X)\big) = \int f(x)d\mu(x)$, where $f(x)$ is the integrand and $\mu(x)$ is a probability measure. With the Monte Carlo integration, we first take i.i.d. samples $X_1, \dots X_n$ from probability measure μ, and then take the sample average $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ as the estimator to $E\big(f(X)\big)$. The law of large number ensures that the estimator will converge to $E\big(f(X)\big)$ as $n \to \infty$.

2.2.3.2 Latin Hypercube Sampling

Suppose there is a random variable $Y = G(\boldsymbol{x})$ defined on the domain C, and we want to find out the overall mean of Y. Sample mean will be a natural estimator of $\bar{Y} = \frac{1}{n}\sum_{i}^{n} G(x_i)$, which would require a set of experimental points $(x_1, x_2, x_3, \dots, x_n)$ from C.

Traditional brute force Monte Carlo uses the simple random sampling for generating these experimental points, i.e. $(x_1, x_2, x_3, \dots, x_n)$ are independently identical samples from the uniform distribution $U(C)$. It is trivial to see that the resulting sample mean is unbiased with variance $\frac{Var(G(\boldsymbol{x}))}{n}$. To improve on the efficiency of traditional methods, McKay, Beckman, and Conover (1979) propose Latin hypercube sampling (LHS) in a paper that is now widely regarded as the first paper on design of experiments. An LHS divides the domain C of each x_i into n strata of equal marginal probability $1/n$, and sample once from each stratum. To be more mathematically concrete, an LHS can be defined in terms of the Latin hypercube design (LHD).

A Latin hypercube design (LHD) with n runs and s input variables, denoted by $LHD(n, s)$, is an $n \times s$ matrix, in which each column is a random permutation of $(1, 2, 3, \dots, n)$. Then an LHS can be generated by the following algorithm (LHSA):

Step 1 Independently take s permutations $\pi_j(1), \dots, \pi_j(n)$ of integers $(1, 2, 3, \dots, n)$ for $j = 1, 2, 3, \dots, s$, i.e. generate an $LHD(n, s)$;

Step 2 Take ns uniform random variables $U_i^j \sim U(0,1)$, $j = 1, 2, 3, \dots, s$, $i = 1, 2, 3, \dots, n$, which are mutually independent. Let $x_i = (x_i^1, x_i^2, \dots, x_i^s)$, where $x_i^j = \frac{\pi_j(i) - U_i^j}{n}, j = 1, 2, 3, \dots, s, i = 1, 2, 3, \dots, n.$

Then $D_n = (\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ is an LHS and is denoted by $LHD(n, s)$.

For instance, for generating an LHS for $n = 8$, $s = 2$, in the first step we generate two permutations of (1,2,3,4,5,6,7,8) as (2,5,1,7,4,8,3,6) and (5,8,3,6,1,4,7,2) to form an $LHD(8,2)$ that is stored in the matrix below on the left. Then we generate $16 = 8 \times 2$ random numbers and store again in the matrix on the right.

$$\begin{bmatrix} 2 & 5 \\ 5 & 8 \\ 1 & 3 \\ 7 & 6 \\ 4 & 1 \\ 8 & 4 \\ 3 & 7 \\ 6 & 2 \end{bmatrix}, \begin{bmatrix} 0.9501 & 0.8214 \\ 0.2311 & 0.4447 \\ 0.6068 & 0.6154 \\ 0.4860 & 0.7919 \\ 0.8913 & 0.9218 \\ 0.7621 & 0.7382 \\ 0.4565 & 0.1763 \\ 0.0185 & 0.4057 \end{bmatrix}.$$

Now an LHS can be generated by the following formula:

$$LHS = \frac{1}{8}\left(\begin{bmatrix} 2 & 5 \\ 5 & 8 \\ 1 & 3 \\ 7 & 6 \\ 4 & 1 \\ 8 & 4 \\ 3 & 7 \\ 6 & 2 \end{bmatrix} - \begin{bmatrix} 0.9501 & 0.8214 \\ 0.2311 & 0.4447 \\ 0.6068 & 0.6154 \\ 0.4860 & 0.7919 \\ 0.8913 & 0.9218 \\ 0.7621 & 0.7382 \\ 0.4565 & 0.1763 \\ 0.0185 & 0.4057 \end{bmatrix}\right) = \begin{bmatrix} 0.1312 & 0.5223 \\ 0.5961 & 0.9444 \\ 0.0491 & 0.2981 \\ 0.8143 & 0.6510 \\ 0.3886 & 0.0098 \\ 0.9047 & 0.4077 \\ 0.3179 & 0.8530 \\ 0.7477 & 0.1993 \end{bmatrix}.$$

Plotting this design in a grid of $64 = 8 \times 8$ cells in Figure 2.3, we can observe that the LHS satisfies that each row and column has one and only one point, and each point is uniformly distributed in the corresponding cell. Note that from $(\pi_1(k), \pi_2(k))$, we can tell $\boldsymbol{x_k}$ is located in which cell and (U_k^1, U_k^2) determines the location of $\boldsymbol{x_k}$ in that cell.

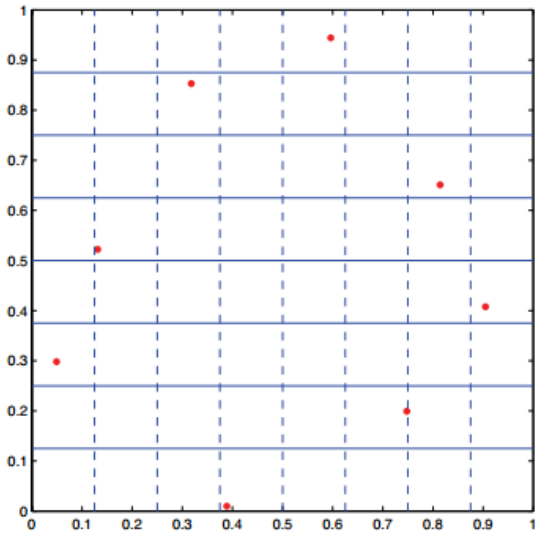


Figure 2.3 An example LHS

In the paper by McKay et al. (1979), they also point out that the LHS has a smaller variance for the sample mean as compared to the simple random sampling. Furthermore, Stein (1987) derives an expression that relates the variance of sample mean in an LHS and simple random sampling:

$$var(\bar{Y}_{LHS}) = \frac{Var(G(\boldsymbol{x}))}{n} - \frac{c}{n} + \circ\left(\frac{1}{n}\right) = var(\bar{Y}) - \frac{c}{n} + \circ\left(\frac{1}{n}\right),$$

where c is a positive constant.

2.2.4 Risk Measures and Risk Preference

Once we obtain the probabilistic distribution for our variable of interest or performance indicator, such as annual energy consumption or the percentage of occupied time that a naturally ventilated building overheats, we need to compute some quantities from such distributions to aid decision making. We refer these quantities as risk measures. For instance, if the main goal of modeling is to compute energy savings compared to the baseline, then the risk measure could be the probability that realized energy consumption is higher than 70% of the baseline amount. With risk measures established, decision makers can explicitly express their risk preferences over them. Taking the above example again, the decision maker may want to limit the risk (probability) that realized energy consumption is higher than 70% of the baseline amount to be below 10%. Although such as statement reads awkwardly, it is much clearer to write a mathematical formula to describe it. Some examples of performance indicators, risk measures and tolerances typical in the building energy domain are given in Table 2.1.

2.2.5 Sensitivity Analysis

Sensitivity analysis has gradually become an intrinsic part of uncertainty analysis for the identification of key input parameters affecting the prediction of building performance. Tian and de Wilde (2011) give a comprehensive overview of the current

state-of-the-art methods. This dissertation will dedicate Chapter 5 on this topic and propose improvements over current methods.

Table 2.1 Example of performance indicators and risk measures

Performance indicator	Typical criterion	Risk measure	Example risk tolerance
Internal rate of return (IRR)	IRR > 5%	Risk (IRR) = Pr (IRR<5%)	Risk (IRR) < 10%
Overheat hours (OH)	OH > 300	Risk (OH) = Pr (OH>300)	Risk (OH) < 10%
Energy reduction compared to baseline (EC)	EC > 0.7 Baseline	Risk (EC) = Pr (EC>0. 7 Baseline)	Risk (EC) < 5%

Note: Pr (IRR<5%) is read as the possibility of project overall IRR is less than 5%.

2.2.6 Iterative Uncertainty Analysis

Uncertainty we obtain from analyzing a pool of buildings other than the one of interest is generally referred as crude or generic. Once the decision-making scenario requires refined data specific to the building of interest, a second iteration uncertainty quantification updates previous knowledge with collected data, for instance, according to a Bayesian framework.

Under the Bayesian framework, both data and parameters can have probability distributions, so we can learn about probabilities of unobservable parameters using Bayes' theorem, which can be written as the following:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)},$$

where $p(\theta)$ is the prior distribution for θ reflecting our uncertainty about the values of θ before observing any data from specialized measurements on this building; $p(\theta|y)$ is the posterior distribution for θ representing our updated uncertainty about θ after taking into account observed data; $p(y|\theta)$ describes the statistical relationship between data and

parameters; the normalizing constant $p(y)$ simply ensures that $p(\theta|y)$ is a valid probability distribution that integrates to 1. Therefore, Bayes' theorem can often be expressed as

$$p(\theta|y) \propto p(y|\theta)p(\theta),$$

where the proportionality is considered with relation to θ.

While $p(y|\theta)$ arises from an assumed sampling distribution for the data, any function of θ, say $L(\theta; y) \propto p(y|\theta)$ will maintain the above proportionality. Consequently, Bayes' theorem essentially states that $posterior \propto likelihood \times prior$. For few simple cases, if we use a conjugate prior distribution for the parameter of interest, the resulting posterior distribution of the parameter is analytically tractable. For some examples of sampling distributions and their corresponding conjugate priors, readers are referred to Table 3.1 in Lunn, Jackson, Best, Thomas, and Spiegelhalter (2012). In more generic cases, where sampling distributions and prior distributions are not conjugate, numerical implementations of Bayesian inference, such as Markov Chain Monte Carlo methods should be used. This dissertation will provide an example using Gibbs sampling in Chapter 4.

CHAPTER 3

UNCERTAINTY TAXONOMY AND REPOSITORY

It is always helpful to build a taxonomy over the topic of research. In this chapter, we present an uncertainty taxonomy, which identifies the relationship between sources of uncertainty, and place model uncertainty in the larger context of the gap between predicted and realized performance. We also document quantified uncertainty sources in the uncertainty quantification (UQ) repository that resulted from a combined effort in the EFRI-SEED project introduced in Chapter 1. The UQ repository is continuously updated.

3.1 Uncertainty Taxonomy

As Kerwin (1993) puts it, there are three types of knowns: unknown knowns, known unknowns, and unknown unknowns. The first kind is also referred to as tacit knowledge, which relates to expertise or intuitions that are frequently used, but difficult to describe or summarize. The second is the major focus of this dissertation - what we know we do not know, or conscious ignorance. The final kind occurs where we do not know what we do not know, which is so-called meta-ignorance. The final kind is the most difficult since we can only learn about it in hindsight.

Given the conceptual understanding of different kinds of uncertainty presented above, we focus mainly on known unknowns in this dissertation, so that we are able to construct a "world view" by defining sources of uncertainty and relating them with each other with enhanced entity relation (EER) diagrams. An EER diagram is a popular tool for systematically describing and defining a subject area of interest. Conceptually, the subject is represented by components (entities) and relationships that define the dependencies or other relationships between entities.

Building simulation is a computational procedure to subject a model form to a defined virtual experiment. We denote the set-up of the virtual experiment as a scenario. Since both the model and scenario may be uncertain, we separate our worldview into two diagrams, describing the physical model (Figure 3.1) and scenario (Figure 3.2).

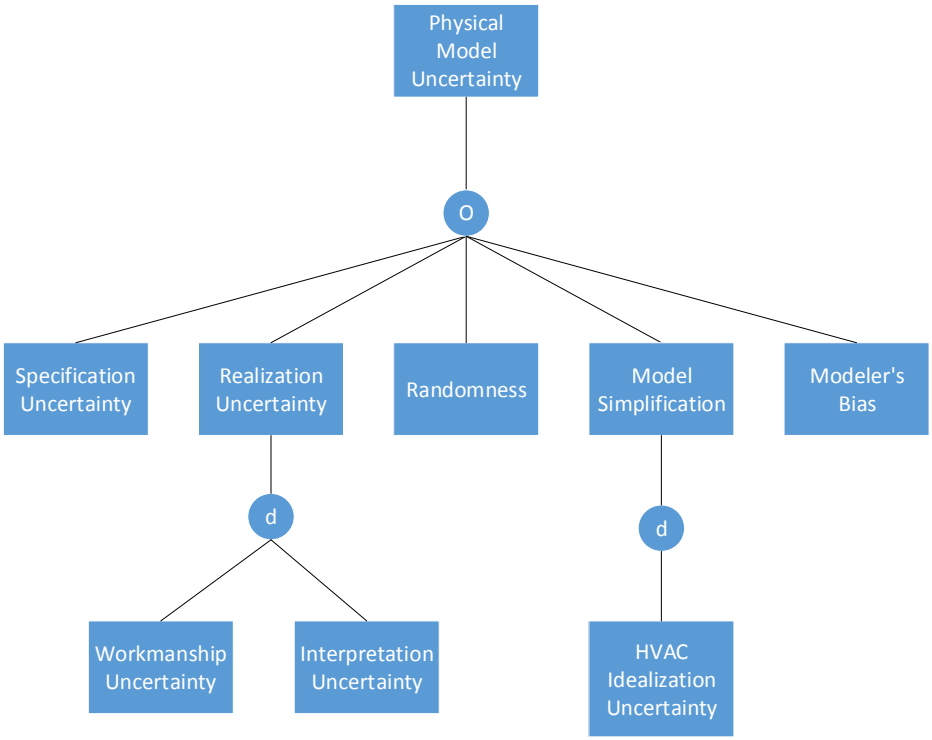


Figure 3.1 Physical model uncertainty EER diagram

The first diagram starts with the overarching term "physical model uncertainty". "All models are wrong, some are useful." (Box & Draper, 1987). When we use models to describe reality with governing relationships or natural laws, imprecision naturally results. Therefore, we refer physical model uncertainty as the bridge between design specs and predictions in design settings. To facilitate understanding, we further break model uncertainty down to specification uncertainty, realization uncertainty, randomness, model simplification and modeler's bias.

Specification uncertainty results from the fact that design documents typically do not fully specify the building of interest at the design stage to the level of granularity and detail required by a simulation program. For instance, when design drawings specify that

one layer of the external walls is made of lightweight concretes, there is still remaining uncertainty around the thermal property of the whole material class of lightweight concretes.

In terms of realization uncertainty, it has to do with the translation from design specifications to a realized building artefact by construction teams. Realization uncertainty could rise from workmanship uncertainty or interpretation uncertainty. The first type of realization uncertainty involves workmanship issues that lead to unintended deviations from the perfect realization, potentially leading to defects and anomalies in a construction detail, such that envelop performance is compromised. The other is more related to vagueness in the construction detailing documents, which leaves room for interpretation by the construction team.

Moving on to randomness, randomness uncertainty in our worldview could potentially be well understood with established theory from other fields but its complexity may not justify its explicit treatment in our simulation. Randomness uncertainty in our domain may subsume known knowns from the manufacturing process, for example, the variability of the thermal property within a batch of bricks may be mainly attributed to the temperature fluctuations and distributions of the kiln, in which they are burned. Since these known knowns are often outside the scope of building simulation, they are only treated as randomness uncertainty.

The fourth physical model uncertainty is model simplification that originates from our idealized representation of reality with models, for instance, the sky is parameterized and divided into three homogenous zones in the modeling of sky diffuse radiation on a tilted surface. Another simplification is control oriented: whereas the actual room temperature fluctuates within a dead-band, this is not typically considered by an energy simulation program, which may assume a perfect controlled room temperature. The latter falls in a subcategory requiring more examination, especially HVAC idealization uncertainty, which rises from the way we simplify and idealize the operation of HVAC

systems. Such uncertainty can be regarded as the aggregate result of many ways in which the internal dynamic interactions between the many different components of an HVAC system and the building can result in deviations from the idealized operation that assumes maintaining perfect temperature set points, delivering the right amount of air into each zone, perfect sensor accuracy, and optimal sequencing of HVAC operation modes.

The last branch of physical model uncertainty is the modeler's bias, which is straightforward, since no two modelers will arrive at the same model given the same building information and simulation tool. The modeler's bias subsumes the combined effects of input parameter inaccuracy as well as the choice of model form (a unique combination of sub-models), for instance, Modeler A may decide to neglect the effect of thermal bridging based on his own experience, relative to the choice of Modeler B. Note that uncertainty associated with a model parameter could represent each of the four categories of uncertainty: specification uncertainty, realization uncertainty, randomness and model simplification, or all combined. One example is the thermal conductivity of a masonry wall made of bricks and insulation layers. The particular type of bricks may not be fully specified by design specs, introducing specification uncertainty, and the inhomogeneous conductivity within a batch of bricks shows randomness. In addition, if we realize that the thermal property of the masonry wall could be compromised by workmanship issues, in particular, by improperly applied insulation layer at wall junctions, realization uncertainty occurs. Finally, by assuming constant conductivity throughout the simulation, we introduce model simplification uncertainty, because in reality, the value should be a function of temperature and moisture content.

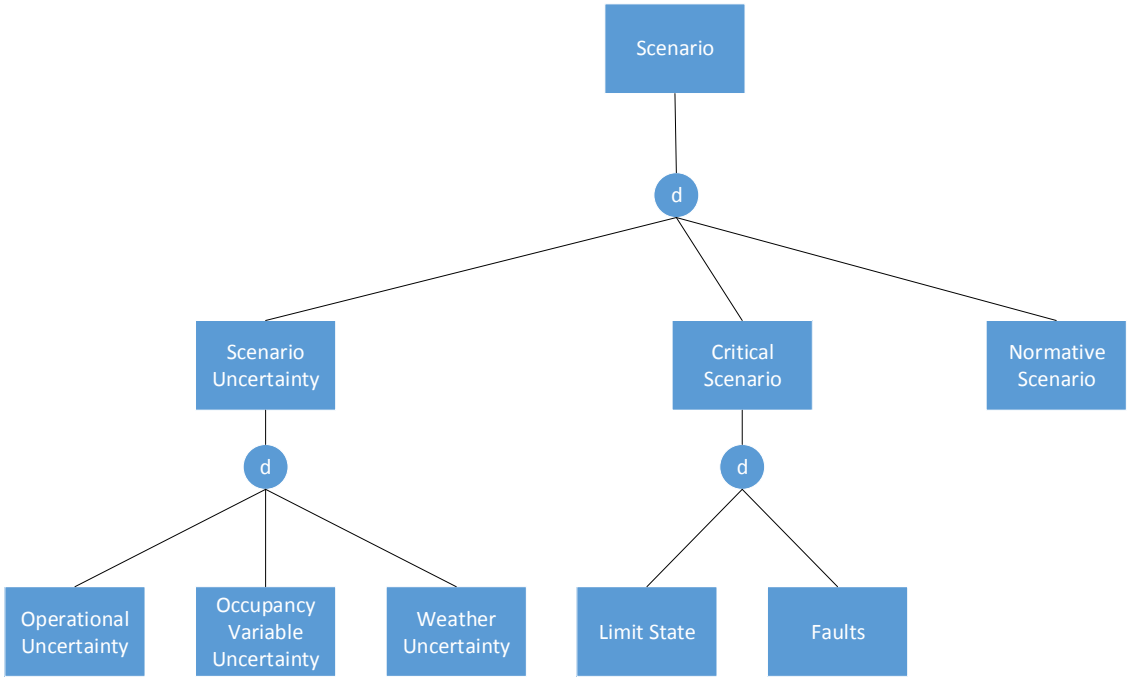


Figure 3.2 Scenario EER diagram

Recall our definition of building performance evaluation, which demands computation of a performance indicator with a specific model subject to a controlled set-up of an experiment. The second diagram deals with the controlled experiment or "scenario" in building simulation studies. Traditionally, the experiment box around building simulation has been deterministic and static. Over the years, the stochastic nature of building operation has drawn more attention and there has been a long lasting debate over stochastic or static scenarios. We see merit in both views and each deserves more research.

The first branch in the taxonomy of scenario presented Figure 3.2 pertains to scenario uncertainty, where the stochastic nature of building exterior and interior environment comes into play. Scenario uncertainty can be further sub-divided into three categories: operational uncertainty, occupancy variable uncertainty, and weather uncertainty. In brief, operational uncertainty relates to operational characteristics of buildings, for instance, the thermostat might experience day-to-day fluctuations and the control logic of HVAC system may not be fully specified in design settings. This is not

be confused with physical specification uncertainty mentioned before: even with the most detailed manual of a particular piece of equipment, lack of full knowledge of their operational schedules still leads to inevitable operational uncertainty. Occupancy variable uncertainty refers to the complex relationship and interaction between occupants and buildings through, for example, their presence and control of lighting and appliances. Weather uncertainty is quite self-explanatory: weather conditions around buildings are also subject to change year to year, especially with the trend of global warming and the effect of urban heat island.

Critical scenarios, on the other hand, describe building operations outside normal conditions. One example is the concept of limit state, borrowed from structural engineering. Limit state is a condition of a structure beyond which it no longer fulfills the relevant design intention. In buildings, limit states include power failure for off-grid solar energy houses, intolerable heat stress in social housing without proper cooling equipment, compromised CO_2 levels due to overcrowding in hospital waiting rooms (Li, Heo, & Augenbroe, 2009), to name a few. This brings up another critical area of deficient building performance where faults are claimed to be one of the major drivers for the performance gap mentioned before. However, we would like to emphasize an important distinction between faults and operational uncertainty. The former denotes system malfunctions due to for instance, lack of proper maintenance or manufacturing defects, while the latter describes variability of building operational parameters under normal conditions, i.e. a system that would pass in regular maintenance inspection.

The last branch in the taxonomy of scenario presented in Figure 3.2 represents another train of thought in building simulation community, which argues that building performance should only be evaluated against normative scenarios. These scenarios are currently most used in ratings systems, where weather, usage, and other uncertainties are eliminated by prescribed values as part of the standard, such that the building itself can be rated without the confounding factor of usage. Such methodology enables a “fair”

comparison among buildings and is adopted in other engineering fields such as the efficiency test of vehicles and refrigerators, where the rating score is reported against standard operating conditions.

Another typology of sources of uncertainty is from Morgan (2009), who suggests a differentiation between (1) uncertainty about the value of empirical quantities that appear as parameters in modeling systems, and (2) uncertainty about the model functional form itself. By summarizing these two cases as parameter uncertainty and model form uncertainty, we then distinguish three sources of uncertainty:

1) model parameters identified as input parameters, that the modeler is supposed to know or choose, but has imperfect knowledge about. We capture this as specification uncertainty and randomness.

2) model parameters that reside inside a sub-model (typically implemented in existing software module) to represent a physical behavior. We are able to capture sub-model deficiency through an uncertainty quantification method and represent this as the range of (some of) the sub-model parameters. This is basically a non-intrusive way of exporting model discrepancy into model parameters. The resulting parameter uncertainties thus act as proxies to model discrepancy. The parameters are not visible to the modeler, but are in fact treated as any other model parameter in the Monte Carlo method.

3) there are no parameters that can subsume the role of (sub-)model discrepancy; this means that there is no parameter available in the model to "expose" a specific model discrepancy. Now we have no other recourse than planting the discrepancy intrusively into the code, i.e. by changing the model form.

In this taxonomy, we define Situation 1 as parameter uncertainty, whereas Situation 2 and 3 are defined as model form uncertainty but implemented in two different forms.

While we contemplate on the gaps that need to be further quantified, it is also helpful to group sources of uncertainty according to their different scales from large to small as shown in Figure 3.3.

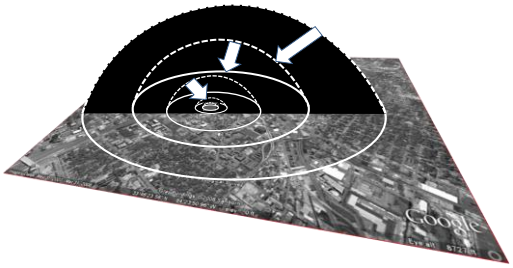


Figure 3.3 Different scales of uncertainty

The first scale is meso-scale. Uncertainties at this scale include stochastic meteorological years, and microclimate variables such as urban heat island, local wind speed and solar irradiation. The next scale is building level, which hosts physical uncertainties such as material properties, convective heat transfer coefficients, infiltration and natural ventilation. The third scale is HVAC system level. Examples are chiller efficiency, duct leakage and system operational uncertainty. The last scale is occupant and process level, since building occupant presence and their control of lighting and appliances are stochastic in nature.

3.2 Uncertainty Quantification Repository

Hereafter we document and refer to quantified uncertainty sources in an uncertainty quantification (UQ) repository from a combined effort in the EFRI-SEED project. We organize them by the scales of uncertainty described above.

3.2.1 Meso-Scale

Uncertainties at this scale include stochastic meteorological years, and microclimate variables such as urban heat island, local wind speed and solar irradiation. The uncertainty of weather can be considered in two ways, in the first approach, each simulation run samples randomly one year of historical weather data, in the other

approach, we study historical weather patterns and generate stochastic representations of them. Then in each simulation run, we sample a different stochastic weather year. The approach for the stochastic modeling of weather can be found in the work by Lee (2014).

Microclimate variables such as urban heat island, local wind speed and solar irradiation are also important boundary conditions for our simulation. Sun, Heo, et al. (2014) quantify the uncertainties in the following most significant microclimate variables: local temperature, wind speed, wind pressure and ground reflectance. In a separate paper, Sun, Heo, et al. (2014) propose improvements over the traditional Perez model (Perez, Stewart, Arbogast, Seals, & Scott, 1986) for the modeling of solar diffuse irradiation on inclined surfaces.

3.2.2 Building Level

At this level, we mainly consider physical uncertainties with material properties, and convective heat transfer coefficients. Exterior convective heat transfer coefficient composes of the effect of natural convection and forced convection. The temperature difference between surface and air influences natural convective heat transfer. As wind velocity goes above 2 m/s, the impact of ΔT on exterior convective heat transfer coefficient becomes very small. Typically, wind forced convective coefficient model relates convective heat transfer coefficients with wind speed in the following form $h = aV + b$, where V is wind speed. Sun (2014) derives the joint distribution of parameters a and b from literature.

Uncertainty of envelope material properties mostly come from Macdonald (2002), which covers most properties of opaque materials and conductivity of glazing. A complete list of material uncertainties can be found in Table 3.1.

Table 3.1 Material property uncertainty

Model parameters	Unit	Source/Reference	Relative standard deviation
Conductivity	W/m-K	Macdonald (2002)	0.05
Density	kg/m3	Macdonald (2002)	0.01
SpecificHeat	J/kg-K	Macdonald (2002)	0.1225
ThermalAbsorptance	-	Macdonald (2002)	0.02
SolarAbsorptance	-	Macdonald (2002)	0.07

3.2.3 HVAC System Level

The current UQ repository lacks detail in this area, but part of this dissertation (Chapter 4 and 6) will zoom into uncertainty with HVAC systems and provide the first steps to adding the missing details. The current repository only considers the effect of manufacturing tolerance for system performance curves, as described in Augenbroe et al. (2013).

3.2.4 Occupant Level

Sun (2014) describes some preliminary work on the quantification of lighting and plug load uncertainty. His work concludes that the peak use of lighting and plug load can be characterized as a normal distribution $N(33.1, 13.1^2)$, when no knowledge about the building of interest is available. However, in practice, modelers have access to certain amount of information on how occupants intend to use the building, either through discussions with designers and future tenants (in case of design) or onsite audits (in case of retrofit). Chapter 4 will propose an updated approach of incorporating modelers' intermediate knowledge on occupancy in probabilistic energy models.

3.2.5 Georgia Tech Uncertainty and Risk Analysis Workbench (GURA-W)

In order to automate the process of uncertainty analysis, we develop the Georgia Tech Uncertainty and Risk Analysis Workbench (GURA-W), details of which are

described in Lee (2014). Figure 3.4 shows the workflow of a typical uncertainty analysis performed by GURA-W.

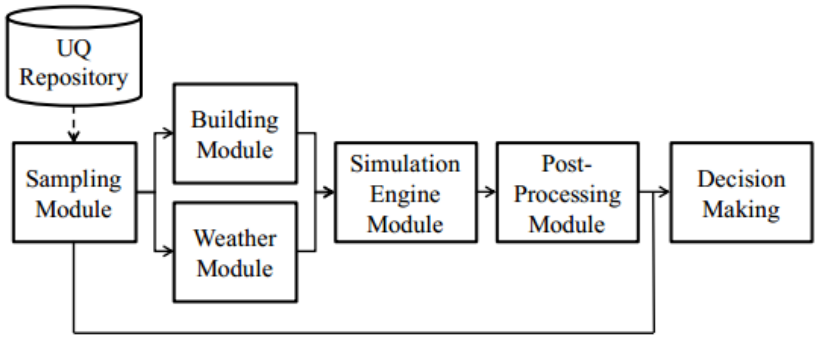


Figure 3.4 Workflow of GURA-W (adapted from Lee (2014))

We briefly describe each of these modules from left to right as follows. The UQ repository is effectively an XML file that stores parameters of the probabilistic distribution for uncertainty variables summarized in this chapter. The sampling module imports default distribution parameters for each uncertain variable (can be overridden afterwards) and then draws samples using Latin hypercube sampling. The weather module either samples randomly from pre-stored stochastic or historical weather years, or only draws the typical meteorological year, depending on user's choice. The building module parses uncertainty variables in the EnergyPlus input files by searching for occurrences of particular identifier tags and exposing these variables for further uncertainty analysis. The connection between sampling module and building module denotes an automated process that feeds Latin hypercube samples of uncertain variables to the simulation engine. The simulation engine is currently EnergyPlus V7.0.0 with modifications that accommodate model form uncertainties such as with microclimate variables. The module specifically calls the RunEPlus.bat batch executable file from EnergyPlus. The post-processing module interprets the output files from EnergyPlus simulations and summarizes the characteristics with variables of interest such as cooling of heating energy consumption. These results will be used in the final decision-making

module, which varies case by case, depending on the decision-making context. Chapter 7 of this dissertation provides such an example.

CHAPTER 4

GAP ANALYSIS AND QUANTIFICATION

With the taxonomy and repository of uncertainty in building simulations identified in Chapter 3, we can readily identify gaps where this dissertation will focus. First, we re-examine the infiltration model (input parameter uncertainty), workmanship issues at a thermal bridge (model form uncertainty), occupancy variables (scenario uncertainty), and HVAC uncertainty (model form uncertainty). The underlying research issues are elaborated in the following sections.

4.1 Infiltration

It is universally acknowledged that infiltration plays a significant role in building energy consumption, especially in terms of heating. Most simulation software adopts the so-called LBL model to represent the outdoor air that enters by way of infiltration. The model is developed by Sherman and Grimsrud (1980) and referred as the "basic" model in the ASHRAE Handbook of Fundamentals (ASHRAE, 2001). The effect of building geometry, leakage area distribution, and terrain and shielding conditions are combined into two reduced parameters in Equation 4.1 such that infiltration can be calculated for any weather condition given the total leakage area of the building. This model targets simple low-rise buildings, which could be simplified into a simple rectangular structure and regarded as a single well-mixed zone.

The basic equation used to calculate infiltration with this model is:

$$Q(t) = \frac{(ELA)\cdot s(t)}{10000}, \tag{4.1}$$

where

$$s(t) = \sqrt{f_s^2 \cdot \Delta T + f_w^2 \cdot v^2} \text{ or } s(t) = \sqrt{C_s \cdot \Delta T + C_w \cdot v^2},$$

$$f_s = \left(\frac{1+\frac{R}{2}}{3}\right)\left(1 - \frac{X^2}{(2-R)^2}\right)^{\frac{3}{2}} \left(\frac{g \cdot B_H}{T_o}\right)^{\frac{1}{2}},$$

$$f_w = \frac{1}{2}\left[(1-R)\frac{\sum_j \sqrt{|C_{p,j} - C_o|} \cdot A_j}{\sum_j A_j} + R\sqrt{|C_o|}\right],$$

$$X = \frac{A_c - A_f}{A_0},$$

$$R = \frac{A_c + A_f}{A_0},$$

$$C_o = \frac{\overline{C_{p,w}}}{1+\left(\frac{A_l}{A_w}\right)^2} + \frac{\overline{C_{p,l}}}{1+\left(\frac{A_w}{A_l}\right)^2},$$

In these equations:

ELA = effective air leakage area [cm^2],

A_c = leakage area in the ceiling plane [m^2],

A_f = leakage area in the floor plane [m^2],

A_0 = total leakage area of the building [m^2],

Q = the air flow [m^3/s],

B_H = height of the building [m],

T_o = internal air temperature [K],

g = acceleration of gravity [9.8m/s^2],

v = local wind speed [m/s],

$C_{p,j}$ = surface average wind pressure coefficient of orientation j,

A_j = surface area of orientation j,

C_o = internal pressure coefficient,

$\overline{C_{p,w}}$ = average wind pressure coefficient for windward surfaces,

$\overline{C_{p,l}}$ = average wind pressure coefficient for leeward surfaces,

A_w = windward effective leakage area,

A_l = leeward effective leakage area.

Caution should be paid with the interpretation of the A_c and A_f: the "floor" leakage is defined as a leakage site that is located at (or near) the level of the building floor that rests on the basement walls, slab-on-grade or crawlspace; the "ceiling" leakage is a leakage site that is at (or near) the ceiling level of the upper story of the building. However, due to the availability of on-site test data, modelers are usually quite ignorant of the building leakage characteristics such that “engineering judgment” and rule of thumbs are frequently followed. As a consequence, generic uncertainty in regard to parameters ELA, R and X exists where no building particular blower door test result is available. Besides, the translation of measured wind speed at a meteorological station to local wind speed and the wind pressure coefficient on façades influenced by urban contexts create additional uncertainty. They are treated in Sun, Heo, et al. (2014), where in principle, the difference between the output of a simple model and a higher fidelity model for the prediction of wind speed/wind pressure is represented by a statistical model parameterized by measurement height z and the wind incident angle. The following will mainly describe the uncertainty quantification of parameters ELA, R and X.

4.1.1 Air Leakage Area

Air leakage is difficult to estimate. In order to obtain data on how leaky a building is, a blower-door test could be deployed when a large blower mounted in a door introduces a uniform pressure difference across the building envelope, while the airflow required for maintaining the pressure is recorded. The calculation of the effective air leakage area is governed by the following equation:

$$Q_f = \frac{ELA}{10000}\sqrt{\frac{2\Delta p_r}{\rho}}\left(\frac{\Delta p_f}{\Delta p_r}\right)^n, \tag{4.2}$$

where

Q_f = measured flow rate at Δp_f [m^3/s],

ρ = air density [kg/m^3],

Δp_r = reference pressure difference [Pa],

Δp_f = applied pressure difference [Pa],

n = pressure exponent [dimensionless].

A reference pressure difference of 4 Pa is frequently used since it represents best the pressure difference under real life conditions.

We analyze measured whole-building envelop airtightness data summarized by Emmerich and Persily (1998). The majority of these data are previously reviewed by Persily (1998) with some additional buildings added. The air leakage values are normalized by the area of the aboveground portion of the building envelop. A preliminary exploration of the data (Figure 4.1) shows the probability density function of measured leakage flow rates ($m^3/s.m^2$) at reference pressure difference of 75 Pa.

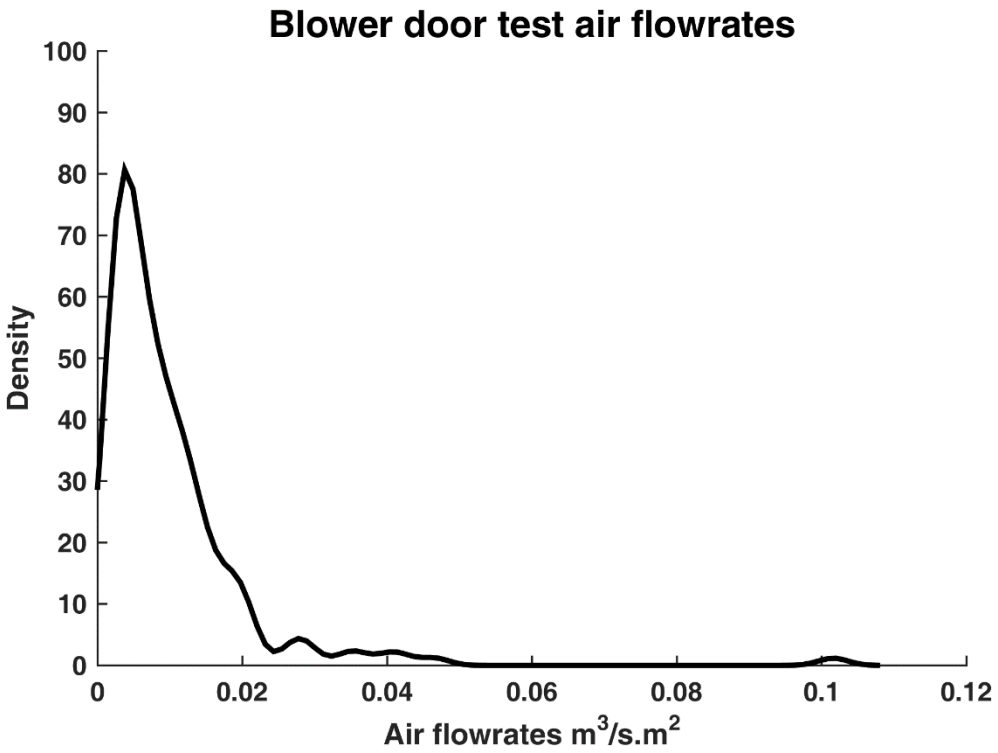


Figure 4.1 Leakage flow rates ($m^3/s.m^2$) at reference pressure difference of 75 Pa

By taking a log transformation of Equation 4.2, we derive:

$$Log(Q_f) = Log(ELA) + \beta_0 + \beta_1 * n - Log(10000), \quad (4.3)$$

where β_0 and β_1 = constants after taking log transformation of Equation 4.2

In Equation 4.2, n, the pressure exponent is typically derived from a regression analysis of blower door test results. n, depending on the characteristic shape of the orifice, is generally difficult to estimate and therefore uncertain. Fortunately, in the building leakage domain, 0.65 for the flow exponent is a good estimate with some variability around it as shown below in Table 4.1. A kernel density fit (Figure 4.2) provides a better visualization of n's distribution. The uncertainty of n will propagate, through Equation 4.3, into the uncertainty of the ELA for each building.

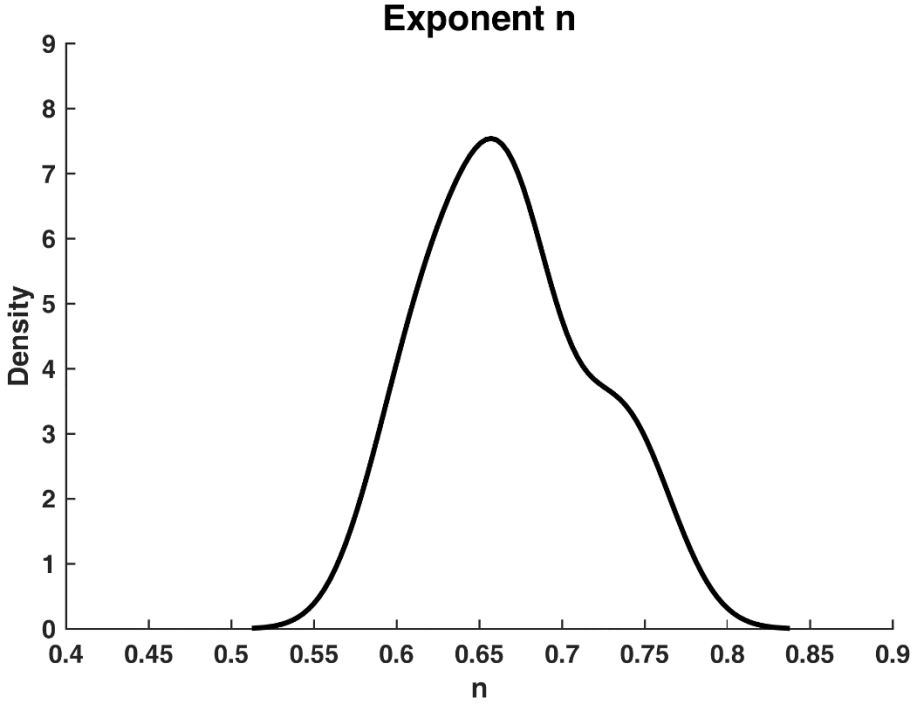


Figure 4.2 Kernel density estimate of the pressure exponent n

Table 4.1 Flow exponent from the literature

Case		n	Reference
House 01	fit to measurements	0.72	W. M. Wang, Beausoleil-Morrison, and Reardon (2009)
House 02	fit to measurements	0.67	
House 04	fit to measurements	0.65	
House 05	fit to measurements	0.68	
House 06	fit to measurements	0.64	
House 07	fit to measurements	0.64	
House 08	fit to measurements	0.59	
House 09	fit to measurements	0.67	
House 10	fit to measurements	0.66	
House 12	fit to measurements	0.63	
House 13	fit to measurements	0.62	
House 14	fit to measurements	0.68	
House 17	fit to measurements	0.67	
House 18	fit to measurements	0.69	
House 19	fit to measurements	0.61	
House 21	fit to measurements	0.62	
Masonry house (pressurization)	fit to measurements	0.76	Walker and Wilson (1993)
Masonry house (depressurization)	fit to measurements	0.74	
Reference house (pressurization)	fit to measurements	0.74	
Reference house (depressurization)	fit to measurements	0.66	
Case	fit to measurements	0.60	Deru and Burns (2003)
Average in cases	fit to measurements	0.73	Jokisalo, Kurnitski, Korpi, Kalamees, and Vinha (2009)

Finally, in order to estimate the ELA for an unknown building, we adopted a hierarchical approach because we tend to recognize that it is unlikely that all buildings have the same underlying ELA, but we also tend to assume that knowing something about other buildings tell us at least something about the one of interest. For example, suppose (the ELA of each building) $\theta_i \sim N(\mu, \sigma^2)$, $i = 1, \dots, N$. The data y_i are a group of observations for each building and they are assumed to be conditioned on the corresponding building-specific parameter θ_i. This is illustrated in Figure 4.3 below. In this hierarchical model, we learn about θ_i, the ELA distribution of one building, not only through direct information from y_i, but also through indirect information which comes from data of the remaining buildings via the population distribution, which is parameterized by ϕ. In other words, parameter ϕ defines the population distribution at a higher hierarchy.

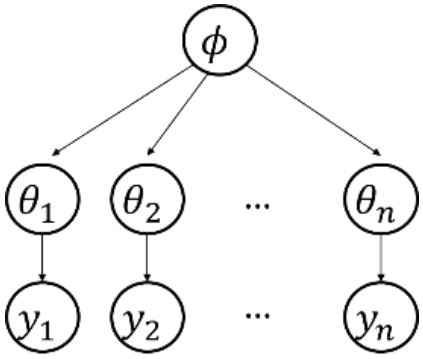


Figure 4.3 A hierarchical model

We implemented this Bayesian technique with BUGS (Bayesian inference Using Gibbs Sampling) with the following steps:

1) For each building in the dataset summarized by Emmerich and Persily (1998), obtain multiple samples of ELA by sampling n;

2) For building i and ELA sample j, assume:

$$Log(ELA)_{ij} \sim N(\mu_i, {\sigma_1}^2) ;$$

3) On a higher level, assume:

$$\mu_i \sim N(\mu, \sigma^2)\ ;$$

4) Assign non-informative priors on μ_i, ${\sigma_1}^2$, μ and σ^2.

With the co-simulation of Matlab and Openbugs, we obtain Gibbs samples of several variables of interest, especially the posterior predictive distributions of $Log(ELA)$ of a particular building and the population mean air leakage ($m^3/h/m^2$) at 75 Pa. Figure 4.4 shows that the ELA of a future building is likely to follow a log-normal distribution with mean 1.282, standard deviation 0.879.

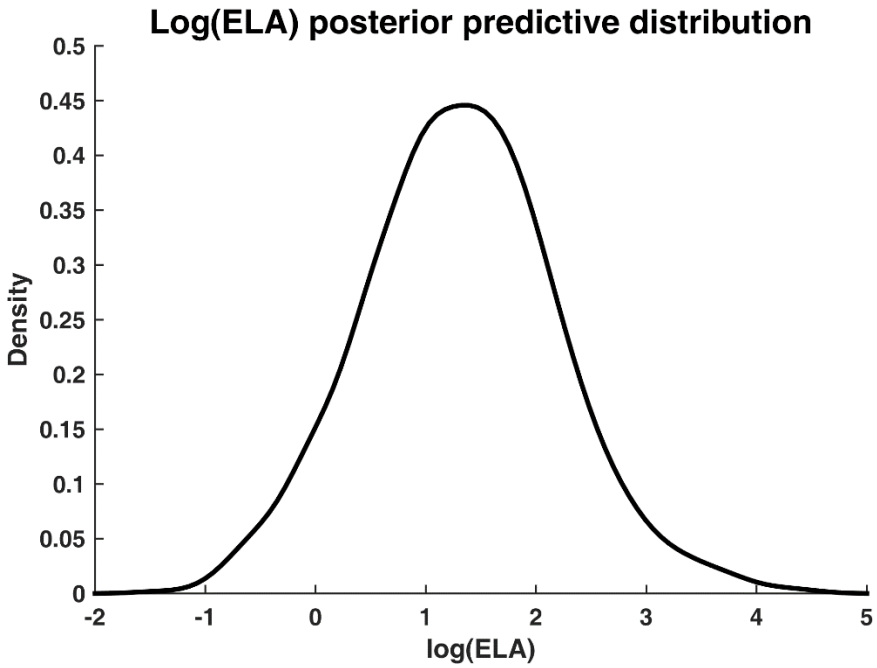


Figure 4.4 Posterior distribution of predicted log(ELA)

We are interested in the population mean air leakage ($m^3/h.m^2$) at 75 Pa (Figure 4.5) because it indicates on average how leaky buildings are and enables us to cross compare with results from other publications. For instance, in the study of Chan, Nazaroff, Price, Sohn, and Gadgil (2005), they show the mean leakage of all buildings in the database is 28.4, which falls around the upper limit of our predicted distribution.

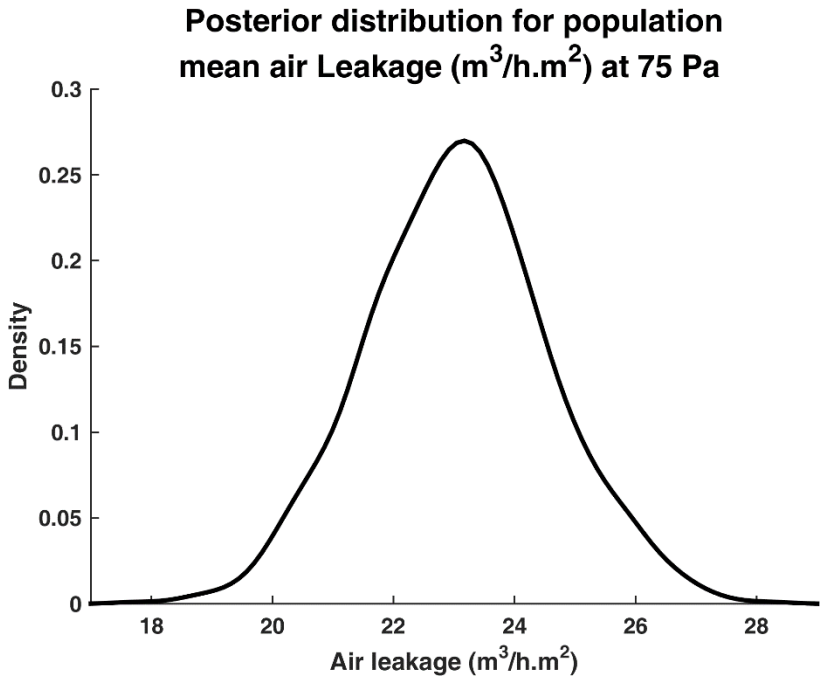


Figure 4.5 Posterior distribution for the population mean air leakage ($m^3/h.m^2$) at 75 Pa

4.1.2 Air Leakage Distribution

One of the critiques of LBL model is that in an effort to simplify the prediction of air infiltration, the model input requires an estimation of the air leakage distribution of the building, which is very difficult to obtain even by measurement. Fortunately, Reinhold and Sonderegger (1983) reports that model predictions are only weekly sensitive (0 – 15%) to the leakage distribution, namely X and R for average houses, thus a very precise determination of them may not be cost effective.

A search through the literature reveals pairwise estimations of R and X for different buildings as summarized in Table 4.2.

Table 4.2 R and X in literature

R	X	ceiling	floor	Reference
0.5	0	0.5	0	Walker and Wilson (1998)
0.3	0	0.3	0	
0.1	0	0.1	0	
0.3	0.1	0.2	0.1	
0.5	0.3	0.4	0.1	
0.65	-0.25	0.2	0.45	
0.85	-0.05	0.4	0.45	
0.5	0.1	0.3	0.2	Swan, Ugursal, and Beausoleil-Morrison (2009)
0.35	0.05	0.2	0.15	Purdy and Beausoleil-Morrison (2001)
0.5	0.1	0.3	0.2	
0.52	0.17			Bassett, 1986
0.48	0.12			
0.44	0			
0.64	0.04			
0.46	0.26	0.36	0.1	Jokisalo et al. (2009)
0.7	0.3	0.5	0.2	
0.5	0	0.25	0.25	ASHRAE (2001)
0.33	0.33	0.33	0	
0.43	0.31	0.37	0.06	Kalamees, Korpi, Eskola, Kurnitski, and Vinha (2008)
0.1	0.06	0.08	0.02	
0.56	0.38	0.47	0.09	
0.39	0.29	0.34	0.05	

We propose correlated sampling of these two parameters for a building lack of specific leakage distribution information with the following approach.

1) Obtain mean vector M and covariance matrix A of R and X vectors from data above;

2) Apply Cholesky decomposition to A and obtain a lower triangular matrix U where $A = LL'$;

3) Randomly sample from a standard normal distribution and obtain a vector X whose two components are uncorrelated.

4) The random vector $Z = M + UX$ is the correlated sampling results of R and X that we need.

Ultimately, a discussion on the steady-state value of internal pressure is provided. It is simply a weighted mean of the steady-state windward and leeward pressure coefficients by the leakage areas A_w and A_l. For simplicity, if we assume $A_w = A_l$, the steady-state internal pressure coefficient is simply the arithmetic mean of $\overline{C_{p,w}}$ and $\overline{C_{p,l}}$.

4.2 Workmanship Issues at a Thermal Bridge

A thermal bridge denotes the part of the building envelope with significantly higher local heat flux than other "undisturbed" parts, which compromises the overall thermal resistance of the entire envelope. The severity of a thermal bridge is typically measured by the following quantitative measures.

1) Temperature factor at the internal surface:

$$f_{R_{Si}} = \frac{R_T - R_{si}}{R_T} = \frac{T_{si} - T_e}{T_i - T_e},$$

Where

$f_{R_{si}}$ is the temperature factor,

R_T (m^2·K/W) is the total thermal resistance of the building envelope,

R_{si} is the internal surface resistance,

T_{si} (°C) is the internal surface temperature,

T_i and T_e are indoor and outdoor temperatures.

Temperature factor shows the relation of the total thermal resistance of the building envelope to the thermal resistance of the building envelope without the internal

surface resistance. Clearly, it is dependent on the internal surface resistance (usually reasonably constant over the length of the measurement time) and the thermal conductivities of the building materials.

2) Incidence factor of the thermal bridge:

$$I_{tb} = \frac{Q_{tb}}{Q_{1D}} = \frac{\sum A_n(T_i - T_{si\_n})}{A(T_i - T_{steady})}, \quad (4.4)$$

Where

I_{tb} is the incidence factor,

Q_{tb} (W) is the heat flux in real conditions,

Q_{1D} is the heat flux through the building envelope as if in absence of the thermal bridge; n is the index of finite elements on the construction surface,

A_n (m^2) is the area of a particular finite element n, while A is the total surface area considered,

T_{si_n} is the temperature at the finite element n, T_{steady} is the temperature with clear wall area.

3) Fraction of temperature variation along a linear thermal bridge:

$$\text{Fraction of temperature variation} = \frac{T_{si\_n} - T_{si\_n-1}}{T_{steady} - T_0},$$

Where

T_0 is the temperature at the most troubled spot.

4.2.1 Methodology

Currently there are several ways of modeling thermal bridging in building performance simulation: 1) use the equivalent wall method to calculate an equivalent wall of three homogeneous layers, which have the same dynamic thermal behavior as the thermal bridge (Martin, Escudero, Erkoreka, Flores, & Sala, 2012); 2) first perform a two-dimensional steady-state simulation of the thermal bridges (the most common tools are THERM or KOBRA), and then export the results to a building performance

simulation to model explicitly the influence area within the envelope; 3) rely on catalogue linear thermal transmittance data from ISO 14683 (ISO, 2007) as a first-order approximation. Naturally, the complexity of the three methods decreases as their demands on fidelity decrease, and it is at the discretion of the modeler to choose the most appropriate one for his case. Furthermore, we argue that an uncertainty analysis should not second-guess the modeler's choice for the following reasons: 1) in the case that the modeler uses the equivalent wall method (recommended by EnergyPlus manual (DOE, 2010)), there is little remaining uncertainty; 2) in some construction details, no additional thermal transmittance due to thermal bridging needs to be added, one example being the following intermediate floor construction with a continuous external insulation layer (Figure 4.6).

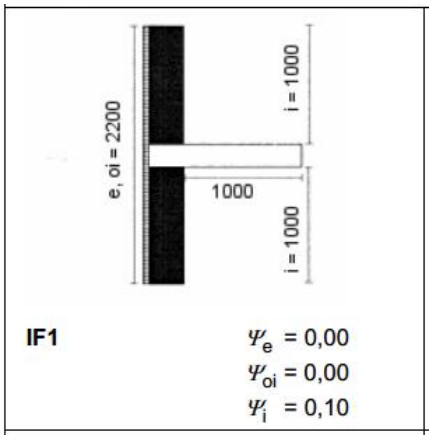


Figure 4.6 A construction detail with no thermal bridging effect (adapted from ISO (2007))

Another layer of complexity comes from construction workmanship that could lead to unintended defects and anomalies in a construction detail. Poor construction detailing design and workmanship issues exacerbate existing thermal bridges. This section engages an initial effort in guiding the modeler through the choice of a reasonable quantification of the impact of poor workmanship at the thermal bridges. Due to the lack of measurement data pertinent to the effect of workmanship variability, we rely on the in situ measurement data by Moon (2005) to quantify the uncertainty associated with the

severity of thermal bridging. His research examines the construction detail at a cavity wall corner as shown below.

Table 4.3 Material property and dimensions used in the hypothetical case

Material	Conductivity (W/m.K)	Thickness (mm)
Brick (outside)	0.6	200
Air gap	0.28	50
Insulation	0.04	50
Gypsum board (inside)	0.2	40

Moon finds that measured temperature factors are significantly lower than the calculated values, suggesting hidden construction defects possibility. His result shows for the above construction detail with a theoretical temperature factor of 0.86, the actual values are between 0.68 and 0.78. The measured corner temperatures between 13.6 ℃ and 15.6 ℃ are also lower than the calculated value of 17.2 ℃, indicating workmanship issues. Unfortunately, Moon's research only concentrates on the most severe part of the thermal bridges, namely the wall corner, without conducting measurements for the entire wall under influence by the thermal bridge. We take the following steps for making further inference on Moon's data:

1) a theoretical Kobra (PHYSIBEL., 2002) model is built for the thermal bridge type present at the test buildings by Moon (2005).

2) we model an artificial heat flux due to workmanship at the wall corner, and calibrate the thermal resistance of this component with measurement data at the most troubled spot published by Moon (2005). The underlying assumption is that the actual workmanship issue matches the type that we assume in the initial model, but only differs in severity.

3) with calibrated models, we can infer the temperature distribution along the thermal bridge and calculate the incidence factor (Equation 4.4), which reflects the additional heat flux through the entire zone of influence.

4.2.2 Results

We first present the simulation results for both the theoretical representation of the thermal bridge without workmanship and a hypothetical unintended damage that happened at a cavity wall exterior corner. In the “bad workmanship” scenario, due to faulty installment, the insulation layer is not consistent at the corner area, where the lowest temperature factor is to be expected. The simulation results are obtained with indoor air temperature at 20 ℃ and outside air temperature at 0 ℃. The detailed configuration and comparison is shown in Figure 4.7.

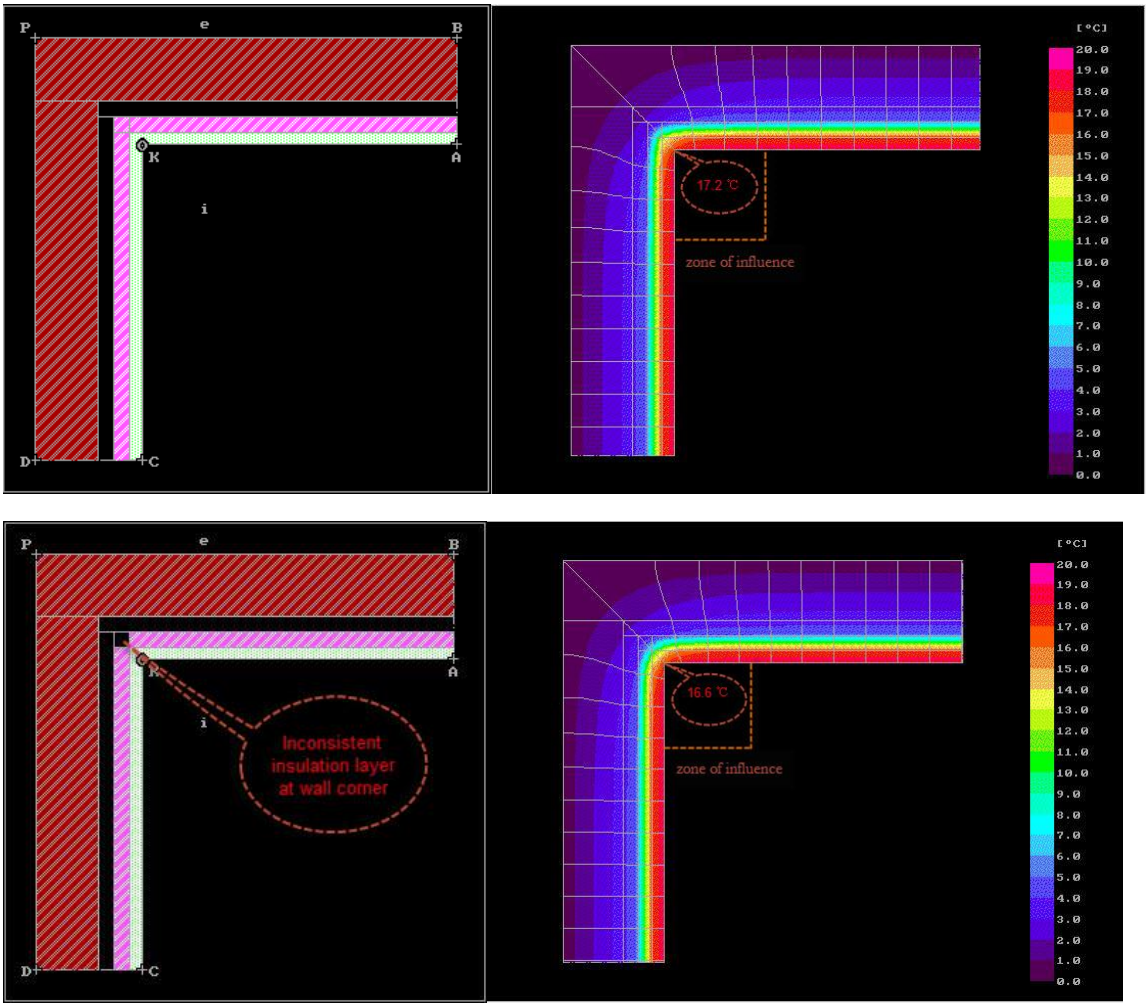


Figure 4.7 “Proper workmanship” (upper) and “Bad workmanship” (lower) scenarios

The zone of influence is determined by examining the heat flow lines in the graphical output and the detailed temperature distribution along the internal wall surface is summarized in the following table.

Table 4.4 Detailed temperature distribution along the internal wall surface

Scenario	Distance from the corner (m)	0	0.02	0.05	0.12	0.19	0.26	0.37	0.63	1
Bad workmanship scenario	Temperature (℃)	16.6	17.8	18.3	18.6	18.7	18.8	18.8	18.8	18.8
	Fraction of temperature variation	NA	55%	23%	14%	4%	5%	0	0	0
Proper workmanship scenario	Temperature	17.2	18.1	18.4	18.7	18.7	18.8	18.8	18.8	18.8
	Fraction of temperature variation	NA	56%	19%	19%	0	6%	0	0	0
Distribution of wall area		Zone of influence						Clear wall		

As we can see from the results in Table 4.4, for the hypothetical faulty workmanship scenario, the temperature of 16.6 ℃ at the most troubled spot at the wall corner is still higher than the measured range between 13.6 ℃ and 15.6 ℃. This indicates the actual workmanship issue is more severe than our assumptions before calibration. In the proposed second step of calibration, we update thermal properties of the modeled air gap, such that the updated temperature sat the wall corner match the lower and upper bounds from measurements, respectively. The inferred in situ temperature distributions are shown below.

Table 4.5 Inferred temperature distributions

Bound	Distance from the corner (m)	0	0.02	0.05	0.12	0.19	0.26	0.37	0.63	1
Lower Bound	Temperature (℃)	13.6	16.5	17.5	18.5	18.5	18.8	18.8	18.8	18.8
Upper Bound	Temperature (℃)	15.6	17.4	18.0	18.6	18.6	18.8	18.8	18.8	18.8

By applying Equation 4.4, the incidence factors of the zone under thermal bridge influence (in this case within 0.26m from the wall corner) are between 1.98 and 1.61, as compared with a theoretical prediction (free from workmanship) value of 1.30. This indicates an additional 23% to 52% of heat transfer due to workmanship for influenced area with regard to the simulation results. We recommend the above identified uncertainty range to modelers who wish to consider workmanship issues at thermal bridges. We suggest modelers first perform a two-dimensional steady-state simulation for thermal bridges and find the influence area within the envelope, and then divide the wall under consideration into two parts: clear wall versus influence area by the thermal bridge. The last step will be to apply the factor of 123% to 152% to the thermal resistance of the influence area, instead of the entire envelop. We also realize more work needs to be done in this area to verify the identified uncertainty range above, preferably with a link to certain construction type, vintages, etc.

More details, including a probabilistic assessment of the impact of workmanship on energy outcomes in office buildings, are given in Wang, Augenbroe, and Sun (2014).

4.3 Occupancy Variables

Occupants participate in the heat balance of the building through their body heat. In addition, occupants' intervention in the control of heating, ventilation and air conditioning systems, and their operation of lighting and appliances have an impact on

building energy consumption. Therefore, the operation and energy use of buildings are highly dependent on the needs and behavior of occupants. In this section, we denote the combination of the temporal and spatial presence, movement and actions of occupants as "occupancy" and variables representing occupancy in building energy models will be articulated in terms of "occupancy variables".

Although building simulation tools continue to evolve and mature, the approach to representing the complex relationships between buildings and occupants has received relatively little attention, and has remained rather rudimentary. Preliminary attempts came up with more realistic and specific diversity profiles that remain invariable from day to day or week to week (Abushakra & Claridge, 2008; Bronson, Hinchey, Haberl, & O'Neal, 1992; Dunn & Knight, 2005; Roberson et al., 2004). Only fairly recently was occupancy recognized as a major obstacle to producing better predictions (Mahdavi, 2010), which has led to a growing body of work that is studying occupancy as stochastic processes, including the IEA-EBC Annex 66 initiative (IEA-EBC Annex 66, 2015). The momentum has, for instance, resulted in tools for generating stochastic demand profiles for occupant services in domestic buildings (Rysanek & Choudhary, 2015). Popular statistical methods used in this type of research are logistic regression analysis, Markov chain models, Poisson process, and survival analysis (Parys, Saelens, & Hens, 2011). For a more detailed survey of related work, readers are referred to Q. Wang, Augenbroe, Kim, and Gu (2016).

The goal of this section is to guide a typical energy modeler through the modeling of occupancy variables when no long-term measurement data about the building of interest are available. In all cases, a typical energy modeler has access to some information relevant to the occupancy of the building that ranges from almost no information to very deep information, as depicted in Figure 4.8. On one end of the spectrum (Point A), we define "total ignorance", which corresponds to the case that the energy modeler has only the most rudimentary information about how the building

intends to be used, which may be the normal case during preliminary design exploration. At the other end of the spectrum, we define “perfect knowledge”, which occurs when the modeler has access to extensive measurement data in terms of occupancy variables. In most situations, one is somewhere between these extremes, which can be characterized as “intermediate knowledge”. In those cases, the modeler has some information about typical building occupancy in part or as a whole, obtained through on-site audits (in case of retrofit) or tenant surveys (in case of new design). However, on-site audits and surveys generally will only convey a limited or even biased expectation of the typical occupancy of the building, while leaving out information about occupancy variability, in other words, the day-to-day fluctuations around the mean. Such information gap could be accounted for and supplemented with detailed data collection and analysis of similar existing buildings, i.e. having the same occupant organization type and employee demographics.

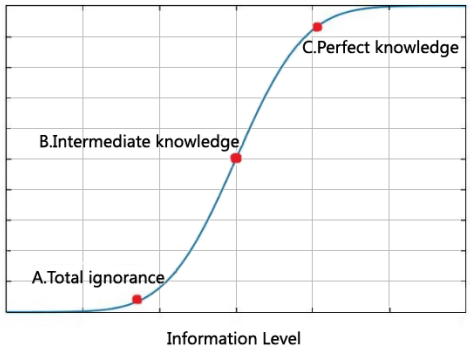


Figure 4.8 Occupancy information available to the modeler (from low to high)

As such, it seems logical to conclude that a desirable modeling framework of occupancy tailored to the above problem statement should incorporate the following features:

- Supporting different strategies of occupancy representation, corresponding to the modeler’s level of knowledge

- Accounting for a potentially biased and uncertain view of the modeler on typical building occupancy
- Being able to extract and learn the characteristics of occupancy variability from a pool of similar buildings and make inference for the building of interest (we thereafter summarize this feature as generalization)
- Consideration of the partial correlation between the variability of occupant presence and actions

4.3.1 Proposed Modeling Framework

Strategies for defining occupancy variables for modelers with ignorant information or perfect information are expected to be straightforward. Therefore, we concentrate on guiding modelers with intermediate knowledge through the creation of an adequate representation of building occupancy. Recall that we have implicitly distinguished between a potentially biased and uncertain view of the modeler on the typical building occupancy, and a general lack of knowledge about the day-to-day fluctuations of occupancy. Such distinction motivates us to model the mean profile and variability of occupancy variables separately with different models.

4.3.1.1 Modeling of Mean Occupancy Profiles

The major assumption is that we do not second-guess an energy modeler with intermediate knowledge about typical building occupancy. In other words, we assume the standard diversity profile has been carefully consolidated with the building owner and is consistent with on-site audits or tenant survey results. This is not to say that these inputs are not subject to inaccuracy. On the contrary, the goal of the model is to generate stochastic samples of mean occupancy profiles based on the modeler's belief about the uncertainty associated with the estimated diversity profiles.

The model of mean occupancy profiles that we propose assumes that the 24-hour profiles follow a multi-variant Gaussian distribution $N(\mu, \Sigma)$ with the value at each hour being a random variable. Since such a distribution is fully specified by the mean vector μ and covariance matrix Σ, we specify μ with the standard input of diversity profiles and diagonal elements of Σ with the variance estimates by the modeler. It is still problematic to generate multi-variant samples since elements of the matrix Σ other than diagonals remain unspecified. These entries determine the correlation between hourly values and they combine to drive the smoothness of the resulting mean occupancy profiles. Our approach to specifying the correlation is for obvious reasons data-driven. We analyze data from ASHRAE Research Project 1093-RP (Abushakra, Sreshthaputra, Haberl, & Claridge, 2001) to establish a correlation matrix C between each hourly random variable and derive the covariance matrix with the following equation:

$$\Sigma = \mathrm{D} * \mathrm{C} * \mathrm{D},$$

where matrix D is a diagonal matrix comprised of the estimate standard deviation of the profile value of each hour. Then a multi-variant Gaussian sample is the following:

$$x = \mu + Lu,$$

where L is the Cholesky decomposition of the positive definite symmetric covariance matrix $\Sigma = LL^T$, and u is a vector composed of multiple independent Gaussian random numbers, each following the standard normal distribution $N(0,1)$. The 24-dimensional sample x represents the mean occupancy of the building of interest.

The above derivation relies on the modeler's estimate of the uncertainty of the standard diversity profiles. Informative estimates can be based on pure judgement, a mixture of data and judgement, or data alone. We provide guidance to modelers with concepts borrowed from Bayesian literature (Lunn et al., 2012). It is well known that in a Bayesian approach, prior information is updated with information from acquired data. In some special cases, the posterior mean is a weighted average of the prior mean and observed data, with the weights determined by the relative contribution of the prior. The

prior information could be interpreted as representing “implicit data”, in other words, the more concrete prior information infers larger “effective prior sample size”. Such is the case in our application: the more reliable the modeler’s information, the less bias and uncertainty with the diversity profile should be expected, that is, higher confidence in the estimates of building occupancy is in some way equivalent to possessing more “implicit data”. However, for a non-existent and unique building, the modeler has to rely on his subjective judgement for an expression of uncertainty. The elicitation of subjective uncertainty distribution can be intimidating and is no easy task. We refer the modelers to O'Hagan et al. (2006) for further reference, as this is not the main focus of this paper. The “mixture of data and judgement” approach will be accounted for in a Bayesian framework that incorporates a tradeoff between subjective beliefs (effective prior samples) and collected actual data samples. Nevertheless, in order to preserve objectivity of this paper, we refrain from expressing our subjective beliefs, but take a data-driven approach by applying a resampling technique called “bootstrapping” to mimic the modeler’s estimate of the standard deviation of the hourly profile values in matrix D. The basic idea of bootstrapping is that the accuracy of summary statistics about a population from sample data (from sample to population) can be inferred by resampling the sample data with replacement (from resample to sample). The simplest bootstrap method starts by computing a resample with replacement of the same size as the original data set. This process is repeated a large number of times, and for each of these bootstrap resamples we calculate a bootstrap mean. We now have a large population of bootstrap means, which leads to an estimate of how much the mean varies. Success of bootstrapping rests on the premise that inference of the empirical distribution of $\hat{J}$, given the resampled data, resembles closely that of the true probability distribution J, given the original data. Note that bias in the original sample data introduces bias to the bootstrapping inference, while variability in the original sample data leads to the uncertainty of the mean. We therefore propose two scenarios to reflect different levels of confidence of the modeler in his

estimate. The first scenario is constructed by choosing two arbitrary days (small effective prior sample size) from monitored occupancy variables as the original sample data and applying the bootstrapping technique to them. It is expected that in this low-confidence case, there will be large bias and variance associated with the mean. The second high-confidence scenario is constructed with a much larger sample size of fifteen days (large effective prior sample size), which will lead to low bias and uncertainty with the bootstrap estimate of the mean.

4.3.1.2 Meta-Analysis of Occupancy Variability

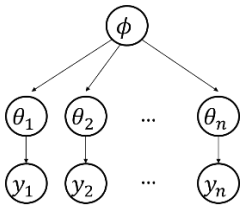


Figure 4.9 A basic hierarchical model

Recall that we assume that a modeler with intermediate knowledge is not aware of the day-to-day variability of occupancy variables, so we need to borrow knowledge from previously thoroughly studied buildings and apply to the one of interest. Recall we have previously defined a general hierarchical data inference structure in Figure 4.3 from Section 4.1. Such a framework also applies here. Suppose data y_i are a group of observations from building i, which we assume to be dependent on the corresponding building parameter θ_i (specific value for each building). We learn about θ_i not only through direct information y_i, but also through indirect information, which comes from the remaining $y_j: j \neq i$, via a higher-level distribution, which applies to the whole population of studied buildings parameterized by ϕ. We therefore propose a meta-analysis that builds upon data from previous research studies, from which we extract and learn the characteristics of occupancy variability and quantify the degree of similarity

between them, eventually enabling the inference of building-specific parameters for the future building of interest.

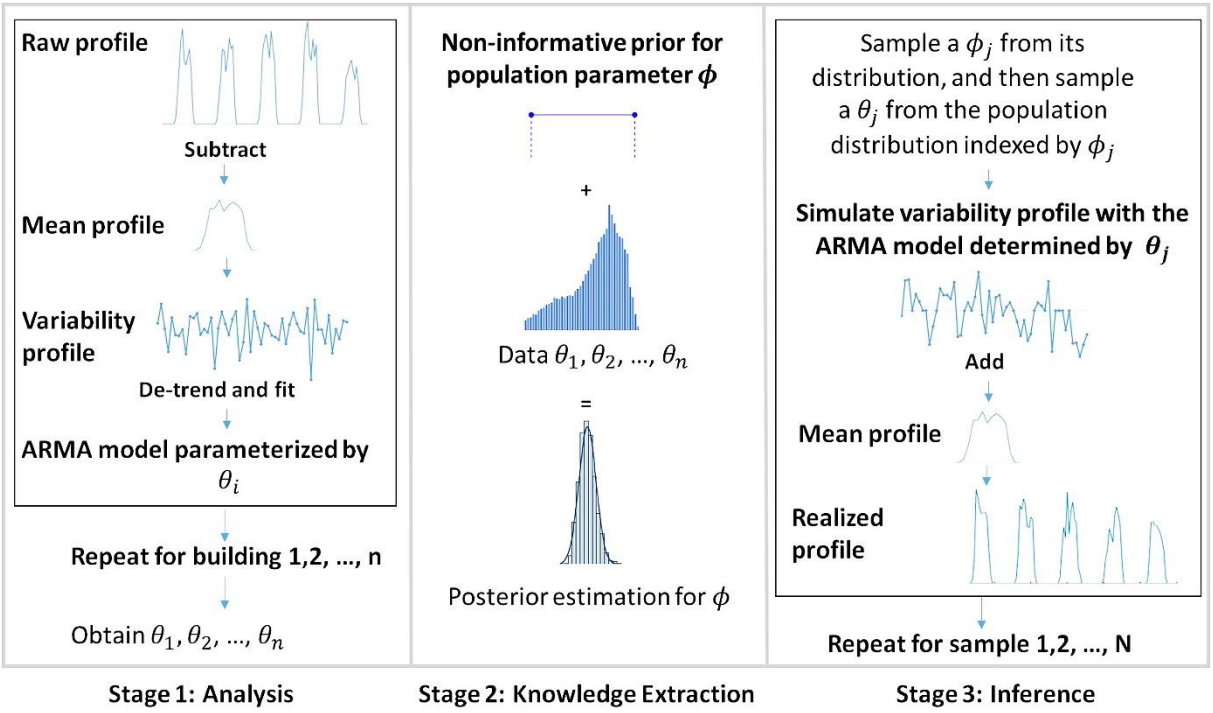


Figure 4.10 Proposed framework for a meta-analysis

We organize the framework into three stages: an analysis stage, a knowledge extraction stage, and an inference stage, as depicted in Figure 4.10. Given collected occupancy data about the number of occupants and lighting/appliance consumption per time step for a total of N weekdays, in the analysis stage, we first normalize these data by the maximum number of occupants and the maximum lighting/appliance consumption from design information, which leads to hourly profiles O_t and L_t normalized between zero and one. Now that we have N samples for each hour of a day for both variables, we derive average occupancy variable values for each hour, which compose a typical weekday. After populating two time series $\overline{O_t}$ and $\overline{L_t}$ with each day being identical, i.e. the constructed typical weekday, we subtract $\overline{O_t}$ and $\overline{L_t}$ from the original series O_t and L_t and obtain the residual series for presence X_t and lighting/appliance usage Y_t, which account for day-to-day occupancy variability. X_t and Y_t are analyzed with linear transfer

function models such that building-specific parameters θ_i are derived. In the knowledge extraction stage, building-specific parameters from multiple buildings are analyzed to estimate the population distribution, parameterized by ϕ. In the final inference stage, for each probable realization of the future building of interest, we perform hierarchical sampling by first sampling a ϕ_j from its posterior distribution, and then sampling a θ_j from the population distribution indexed by ϕ_j. With the building-specific parameter θ_j, we can reproduce occupancy variability with time series models. Having established the general framework, we now move on to more specifics on linear transfer function models.

Serial dependence between consecutive observations, which is common in occupancy data, is called autocorrelation. Autocorrelation is beyond the assumption of independence for most statistical methods, but can be easily taken into account with time series methods. Therefore, the application of time series methods to the study of occupancy seems to be natural. For instance, Schweigler et al. (2009) was able to create a seasonal $ARIMA$ model and robustly forecast a clinic's emergency department (ED) bed utilization 4 and 12 hours in advance for three different EDs. Inspecting the limited research efforts that are available on the topic of occupancy modeling with time series models, we are able to identify the following two obstacles. First, occupancy in general is non-stationary and highly dependent on time of day, so standard $ARIMA$ models cannot be directly applied without data pre-processing, which may be the reason why literature on this topic is scarce. This obstacle is overcome since we concentrate on the modeling of occupancy variability, which is generally stationary and independent of time. The second obstacle is that of the times series models developed around this topic, few allow for the interplay between occupant presence and their operation of lighting and appliances, which has to be examined together with a multi-dimensional time series technique such as a linear transfer function model detailed below.

Consider two time series (X_t, Y_t). X_t is the input time series and Y_t is the output time series. A linear transfer function model has the following form:

$$Y_t = \upsilon(B)X_t + N_t, \tag{4.5}$$

where B is the backshift operator, $\upsilon(B) = \upsilon_0 + \upsilon_1 B + \upsilon_2 B^2 + \cdots$, and N_t is an $ARIMA(p, d, q)$ process. The operator $\upsilon(B)$ is called the transfer function of the model with weights $\upsilon_0, \upsilon_1, \upsilon_2, \ldots$ The system is stable if $\sum_{i=1}^{\infty} |\upsilon_i| < \infty$. In the form of difference equations, Equation 4.5 is written as

$$Y_t = \delta_r(B)^{-1}\omega_s(B)B^b X_t + N_t,$$

where $\delta_r(B) = 1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_r B^r$, $\omega_s(B) = \omega_0 - \omega_1 B - \omega_2 B^2 \ldots - \omega_s B^s$ and b is the delay of b time lags.

Suppose for the moment that $Y_t = \upsilon(B)a_t + N_t$, i.e. $X_t = a_t$, which is white noise independent of N_t. Then it is easy to show that the transfer function is proportional to the cross correlation function of the two series. Now suppose that X_t is an $AR(p)$ process, that is $\phi(B)X_t = a_t$, where $\phi(B)$ is a polynomial. Multiplying both sides of Equation 4.5 by $\phi(B)$, we obtain $\phi(B)Y_t = \upsilon(B)a_t + \phi(B)N_t$. That is for the process $Y_t' = \phi(B)Y_t$, by computing the cross correlation between Y_t' and $a_t = \phi(B)X_t$, we can estimate the transfer function. This is called the pre-whitening of the input. As such, we can summarize the procedures of estimating a linear transfer function model from data:

1) Fit an $ARMA$ model to the input series X_t, retain model coefficients for use in Step 2 and the fitted residuals $\widehat{a_t}$ for use in Step 3;

2) Apply the $ARMA$ operator determined in Step 1 to determine the pre-whitened output series Y_t';

3) Use the cross correlation function between $\widehat{a_t}$ and Y_t' in Step 1 and 2 to suggest an appropriate form for the component of the transfer function $\upsilon(B) = \delta_r(B)^{-1}\omega_s(B)B^b$, especially the estimated time delay b;

4) Obtain model coefficients by fitting a linear regression of the following form:

$$Y_t = \sum_{k=1}^{r} \delta_k Y_{t-k} + \sum_{k=1}^{s} \omega_k X_{t-k-b} + u_t, \quad (4.10)$$

where $u_t = \delta_r(B)N_t$. Retain the residuals $\widehat{u_t}$ for use in Step 5.

5) Apply the moving average transformation $\delta_r(B)^{-1}$ to the residuals $\widehat{u_t}$ to find the noise series N_t, and fit an $ARMA$ model to it.

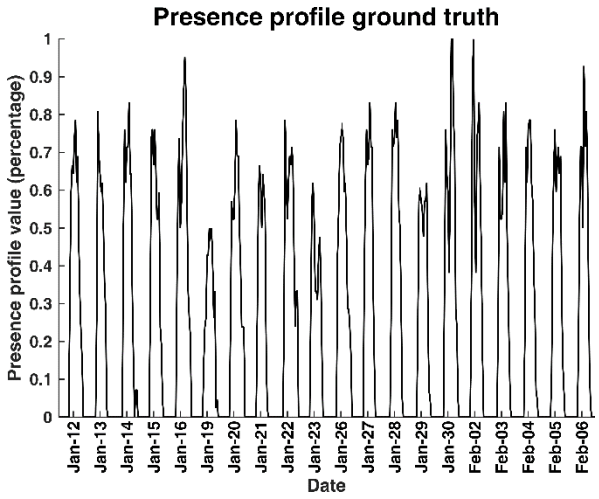
4.3.2 Results and Discussions

Despite the effective framework that we propose above, we realize that the research community possesses only sporadic data sources of building occupancy, especially those that contain detailed interval data regarding both occupant presence and their control of lighting and use of appliances. Therefore, we rely on our own dataset and for now only analyze the characteristics of one building as an illustration of the analysis stage (Stage 1 in Figure 4.10) in the framework. The measurements are conducted from January 11th to February 7th in a medium-size mixed use building on the Georgia Tech campus. The building has an area of about 45000 square foot divided into office and lab spaces. At each entrance of the building, occupancy sensors (in Figure 4.11 below the "exit" sign) based on infrared sensing technology are installed. These directional sensors detect the passage of people, determine their directions, and then send signals to minicomputers (in Figure 4.11 above the "exit" sign), where the signals are transformed into counting data and stored until a download command is received. With these sensors, we derive the total number of occupants inside the building by subtracting the accumulated number of exits from that of entrances at each time step. Data are collected with a five-minute time interval. Furthermore, the building is equipped with electricity sub-meters so that we can collect the lighting and appliance consumption data in five-minute time interval.

Figure 4.11 Occupancy sensor connected with a minicomputer above an entrance

4.3.2.1 Profiles for Occupancy Variables

We exclude weekends from this study since the variability of occupancy is quite small during those days. Figure 4.12 shows the occupancy profiles during twenty weekdays, which happen to show quite large day-to-day fluctuations.



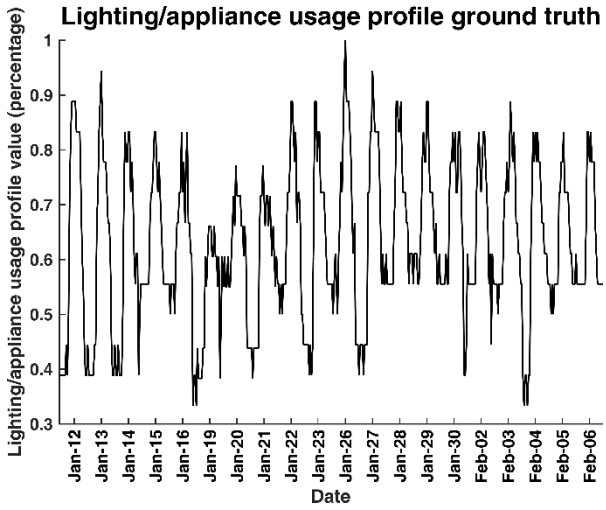


Figure 4.12 Presence profile (left) and lighting/appliance usage profile (right) ground truth

4.3.2.2 Estimation of the Linear Transfer Function Model

Recall that we use linear transfer function models to analyze the variability of occupancy variables, so the profiles in Figure 4.12 have to be further pre-processed. We first subtract the mean occupancy series, remove the non-occupied hours and then further de-trend the series with a linear regression with the week number as a categorical independent variable. Table 4.6 summarizes the statistics for the two regressions. By observing the p value, we identify that in Week 2, both presence and lighting/appliance usage are lower than other weeks. In addition, lighting/appliance usage in Week 3 is higher than other weeks. After this procedure, the residuals of the two series, i.e. X_t and Y_t are stationary, as shown in Figure 4.13.

Table 4.6 Statistics of regression models

Series	Predictor	Coefficients	P Value
Presence	Intercept	0.014	0.39
	Week 2	-0.067	0.003
	Week 3	0.01	0.63
	Week 4	0.0004	0.99
Usage	Intercept	-0.0017	0.83
	Week 2	-0.0035	0.002
	Week 3	0.038	0.001
	Week 4	0.004	0.70

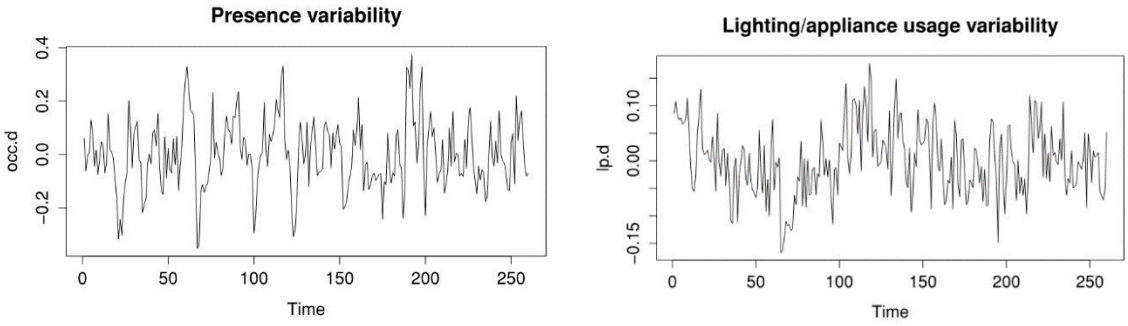


Figure 4.13 De-trended presence variability (left) and lighting/appliance usage variability (right)

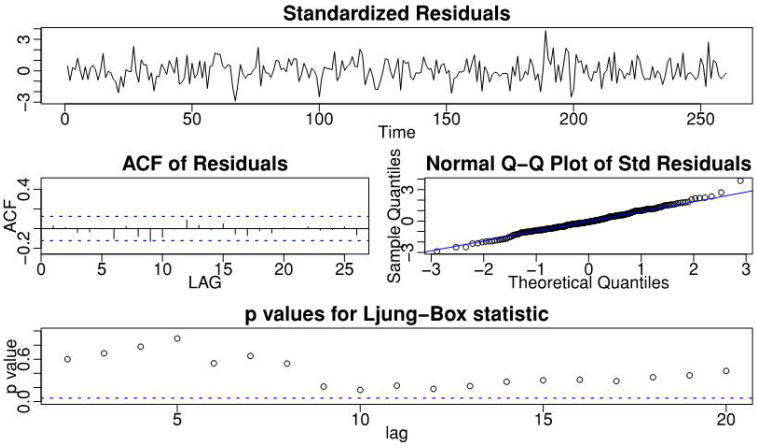


Figure 4.14 Diagnostic plots for the $AR(1)$ model for X_t

Following the procedures for estimating a linear transfer function, we choose an $AR(1)$ model after model selection for the presence variability series X_t: $X_t = 0.59X_{t-1} + a_t$, where $\widehat{\sigma_{a_t}^2} = 0.01082$. Diagnostic plots in Figure 4.14 show that the standardized residuals seem normally distributed, and that auto-correlations of the residuals $\widehat{a_t}$ seem to have been removed, since the Ljung–Box test cannot reject the null hypodissertation that the model is adequate at a 0.05 level. Applying the operator $\phi(B) = (1 - 0.59B)$ to the output series Y_t, we get the pre-whitened series Y_t'. Figure 4.15 depicts the cross-correlation between Y_t' and $\widehat{a_t}$, with an apparent shift of $b = 1$ hour. Through trial and error, it seems plausible to hypothesize a model of the form $v(B) = (1 - \delta_1 B)^{-1} B^1$, and the corresponding regression estimate is

$$Y_t = 0.46Y_{t-1} + 0.09X_{t-1} + u_t,$$

where $\widehat{\sigma_{u_t}^2} = 0.0029$. It also turns out $\widehat{u_t}$ is basically white noise so that the final Step 5 is not necessary. Note that in this particular case, the variability of lighting/appliance usage is only related to usage and presence in the previous hour, but this may not always be the case.

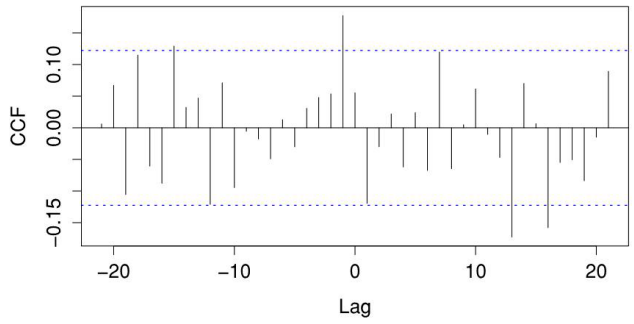


Figure 4.15 Cross-correlation between Y_t' and $\widehat{a_t}$

Comparing the scale of $\widehat{\sigma_{a_t}^2}$ and $\widehat{\sigma_{u_t}^2}$, we can see that the variability associated with presence is much higher than that with usage, which provides a side evidence that, for instance, occupants are found to be less likely to turn off their task lighting in case of temporary absence (Reinhart, 2004).

4.3.2.3 Outcomes of the Model of Mean Profiles and the Variability Model

As an illustration, we show the comparison of mean occupancy profiles in Figure 4.16 and realized stochastic occupancy on a typical weekday in Figure 4.17.

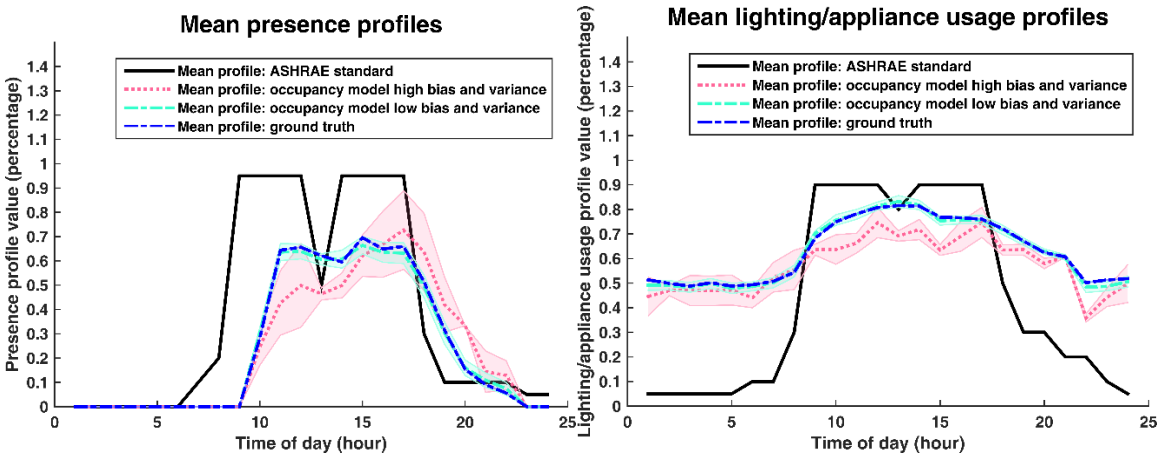


Figure 4.16 Mean occupancy profiles: presence profile (left) and lighting/appliance usage profile (right)

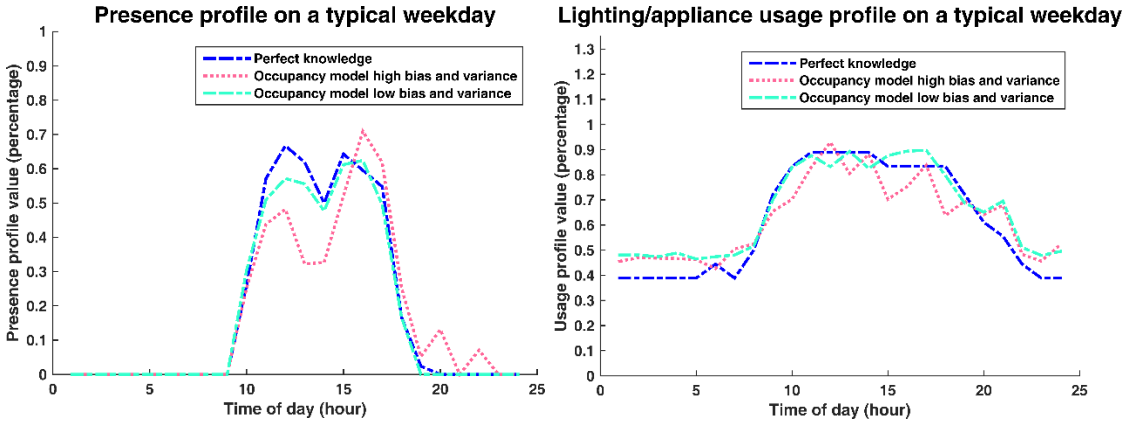


Figure 4.17 Presence profile (left) and lighting/appliance usage profile (right) on a typical weekday

Figure 4.16 highlights the large discrepancy between the ASHRAE standard profiles and mean profiles from the ground truth: the former greatly over-estimates the presence during the day while under-estimates the lighting/appliance usage during non-occupied hours. To be more specific, ASHRAE profiles assume 9.2 person-hours a day and 10.4 lighting/appliance-hours a day, while the ground truth shows 5.9 person-hours a day and 15.1 lighting/appliance-hours a day. This indicates ignorant profiles for presence

and lighting/appliance usage add up to 55% higher and 31% lower levels than the ones from the ground truth, respectively. Figure 4.16 also features a comparison between the two scenarios for the model of mean occupancy profiles, each with an added uncertainty band (reflecting the modeler's confidence in his estimate) around the mean. Indeed, profiles from the high bias and variance scenario deviate significantly from mean profiles from the ground truth, while those from the low bias and variance scenario closely follow the trend of actual occupancy (the green and blue curve virtually overlap). Furthermore, Figure 4.17 shows that a random realization generated by the low bias and variance scenario combined with the occupancy variability model can reproduce the characteristics of actual occupancy variables, while the other scenario (high bias and variance) introduces more inaccuracy.

4.3.3 Relevance of Stochastic Occupancy Models

The imminent question to be answered then is whether occupancy modeling adds significantly to an accurate prediction of building performance. Various researches have reported the effect of occupancy modeling on building energy outcomes (Bonte, Thellier, & Lartigue, 2014; Duarte, Budwig, & Van Den Wymelenberg, 2015; Saelens, Parys, & Baetens, 2011). In particular, recent work by Tahmasebi and Mahdavi (2015) evaluates the impact of different occupancy representation on building performance indicators. However, few among these have explicitly made the important distinction between the impact of occupancy variables on direct and indirect energy outcomes. It is an increasing trend in commercial buildings that control is left to automated systems, leaving little to human intervention, except for direct electricity use for lighting (in some cases) and appliances (in most cases). Another issue to consider is the level of interaction between the operation of building systems and the spatial and temporal variability of occupancy. If controllability of the HVAC system is distributed and nimble, one can expect a strong dependency between the building operation and occupancy (Goyal, Barooah, &

Middelkoop, 2015; Oldewurtel, Sturzenegger, & Morani, 2013). However, the current market is still dominated by conventional HVAC systems with a central and fixed control and limited response to zonal and temporal variability. Adjusting outdoor air supply, controlling zonal temperature or shutting down unused zones based on occupancy, or controlling operable windows on a zone-by-zone basis is typically not possible. A direct response of the building operation to occupancy is mostly absent or at best "muted" in conventional buildings. Therefore, we argue that in most current buildings, the occupants' impact on HVAC system consumption is only the indirect effect of their presence and control of lighting and appliances. Another important consideration is that the impact of occupancy models cannot be examined without acknowledging the role of all other sources of uncertainty. Rather, the sensitivity of occupancy models should be ranked against other uncertainties.

Readers are referred to Wang et al. (2016) for a detailed examination of the relevance of occupancy models on various energy outcomes of interest under uncertainty, with a clear distinction between direct and indirect energy consumption.

4.4 HVAC System Uncertainty

In large commercial buildings, Heating, ventilation and air conditioning (HVAC) systems always represent the largest primary energy end use (Huang, Wang, Huang, & Augenbroe, 2015). Therefore, the importance of truthfully representing HVAC systems in building performance simulation cannot be understated. Conventionally, more attention has focused on the characterization of uncertainty in building thermal parameters, rather than on HVAC system parameters. Various factors such as duct leakage, improper equipment sizing, imprecise control, ageing and degradation, sensor drift, and damper and valve faults are difficult to find out without field inspection, but all potentially contribute to the discrepancy between predicted and observed system

consumption. In addition, our models of HVAC systems represent ideal behaviors, whereas in reality the system may behave astray from idealization.

The reasons why predicted system consumption deviates from what is observed in reality could be loosely attributed to three main causes: 1) the operational schedules and parameters are not known precisely, for instance, the operational hours and thermostat settings are subject to uncertainty, which is characterized under scenario uncertainty in our overall uncertainty taxonomy; 2) our models are associated with model simplification and idealization in the representation of the behavior of HVAC components in reality; 3) HVAC system components are likely to deteriorate or subject to faulty operations over their life span. In this dissertation, we propose to treat the above three issues in different approaches, realizing the scope of this work is prediction under design settings. First, scenario uncertainty can be quantified or eliminated through discussion and planning with the design team and potential tenants of the building, where the same philosophy in the modeling of occupancy variables in the previous section applies. We then propose a bottom-up approach to deal with model idealization and a top-down approach to infer the remaining variability. The rationale behind such a choice is the following. First, the scope of this work is not meant to capture the whole spectrum of performance of the HVAC system during its life cycle, when faults are almost inevitable depending on the maintenance practice. The design prediction model we are aiming for will only serve as a baseline for building commissioning, and help maintenance personnel analyze and identify areas where further inspection needs to focus on for fault detection. Consequently, it is not recommended practice to intertwine all possible HVAC faults with other uncertainties at the design stage, because for one thing, the possible faulty scenarios are too numerous to enumerate, and for the other, taking them into account will exaggerate the uncertainty associated with predictions so much that comparative decision-making scenarios will drown in the results of potential HVAC faults. A bottom-up approach, as the name indicates, starts from an investigation of compositional sub-

systems to gain insight into the overall, inherently complex system. In our application, the bottom-up approach quantifies model form uncertainty with each HVAC component while the Monte-Carlo simulation of the HVAC system will combine the UQ results of all HVAC components into the overall outcomes of the interacting building and HVAC system. It should be noted that this approach relies on the embedded representation of the HVAC model, assuming that the model captures the interaction dynamics well enough so that the propagation of the uncertainties in the individual HVAC components combine into the uncertain behavior of the integrated system. This effort requires the joint collaboration of the building and HVAC research community. This section will present an example to demonstrate the approach. It targets the quantification of uncertainty in the model of an air handing unit with VAV terminals in EnergyPlus to demonstrate the methodology. We postulate that after eliminating or quantifying all other sources of uncertainty and explicitly modeling inaccuracy introduced by HVAC model simplification, our predictions will match consumption of the building right after commissioning, when the newly built building is free from system faults, degradation, etc. For an existing building for which we have access to measurement data, rudimentary calibration to an as-operated model will be necessary. This would make little sense without the recognition of potential faults, adding another layer of complexity to the bottom-up approach. In such case, it seems more advantageous to switch to a top-down approach which lumps all remaining or additional HVAC uncertainties (which could be minimal if we have done a good enough job in the bottom-up stage) and the impact of faults (if any) into an overall HVAC "uncertainty factor". To elaborate, the top-down approach in this dissertation is based on six selected campus buildings, which have been, modeled as truthfully as possible in terms of all input parameters, model form and scenario uncertainties. They are subjected to an UA and the resulting probabilistic predictions are compared against measurement data. We then superimpose an HVAC uncertainty factor onto the probabilistic predictions that would lead to a close match

between UA outcomes and measurement. This is accomplished with a parametric analysis in Chapter 6, until the statistical measure of the goodness of the model is passed. The resulting uncertainty factor can be interpreted as mainly caused by the ensemble of unquantified sources attributed to HVAC idealization and system faults that are likely to exist to some (unknown) extent in the buildings we are investigating. The remainder of this section describes the bottom-up approach, while a separate section in Chapter 6 will be dedicated to the top-down approach.

4.4.1 Literature Review

From our literature review in Chapter 1, we realize that most of previous uncertainty analysis research focuses on climate conditions, and building architectural parameters, such as envelope properties, and internal loads. Few studies investigate the impact of HVAC system representation on building energy predictions. Sun, Gu, Wu, and Augenbroe (2014) propose a framework for HVAC system sizing, taking into account uncertainty associated with building thermal loads. Augenbroe et al. (2013) investigate whether the practice of modeling the HVAC system as an integral part of the building simulation is justifiable, in other words, whether the fully coupled dynamic approach could be replaced with an uncoupled sequential approach (i.e., using a dynamic building simulation followed by a simplified HVAC calculation). In their study, building parameter uncertainties as well as families of equipment performance curves reflecting manufacturing tolerances are incorporated. Rasouli, Ge, Simonson, and Besant (2013) evaluate the effectiveness and economic performance of energy recovery ventilators (ERVs) under arbitrary uncertainty associated with building and HVAC system input parameters. Wang et al. (2012) study the combined effect of various building operation practice, such as HVAC equipment schedule, VAV box minimum flow setting, and room temperature set-points for occupied hours. However, most input parameters in their study are associated with parametric variations instead of rigorously quantified uncertainty

range, for instance, minimum airflow fractions of 15%, 30%, and 50% of the maximum airflow are used to represent the range of practice. Another study by Wang (2014) identifies a sampled list of maintenance issues, including cooling tower fouling, boiler/chiller fouling, refrigerant over or under charge, and temperature sensor offset, etc., and study their energy impact for an office building. Huang, Wang, et al. (2015) propose a method of estimating the aging effect of chillers with a Bayesian Markov Chain Monte Carlo method, which is probabilistic in nature compared to a conventional deterministic maintenance factor approach (0.01 for systems that undergo annual professional maintenance, and 0.02 for those that are seldom maintained). Yan (2016) analyses an HVACsim+ model along with field measurement data and identifies the inadequacy of the EnergyPlus model in representing one type of minimum outdoor air control in VAV systems with a fixed minimum damper position to ensure minimum outdoor air supply. The results show that when the damper is at a fixed minimum position, neither the outdoor airflow rate nor the outdoor air ratio is a fixed value in actual systems, contrary to the model assumptions in most building energy simulation models, such as EnergyPlus.

In particular, this section wishes to investigate the potential model inadequacy of the simplified representation of VAV systems in EnergyPlus. This type of system is chosen for our investigation because of its wide application. Various research has investigated into VAV systems from different perspectives. Aynur, Hwang, and Radermacher (2009b) simulate a VAV system in an existing office building with EnergyPlus with carefully prepared input parameter values that match the reality, and they find 71% simulated power consumption data fall within ±15% range from the measured data. In another study by Aynur, Hwang, and Radermacher (2009a), they compare VAV and variable refrigerant flow (VRF) systems in an existing office building environment and find that overall VRF systems deliver 27.1% to 57.9% energy savings in regard to VAV systems, depending on system configuration, indoor and outdoor

conditions. Mei and Levermore (2002) fit a polynomial curve and artificial neural network model to represent the non-linear characteristics of a fan, and inset the fan and terminal box models into HVACSIM+. Yoshida, Kumar, and Morita (2001) propose recursive autoregressive exogenous algorithm as a dynamic fault detection model for VAV air handling units, and their results are validated with data obtained from a real building after introducing artificial faults. Yang, Jin, Du, Fan, and Chai (2011) compare four control strategies of VAV systems in terms of indoor air temperature, energy consumption, CO_2 concentration and predicted mean vote in an office building in China. Wang and Jin (2000) propose an online model-based methodology for the optimal control of VAV systems. Looking through the literature, insufficient attention has been paid to the investigation of the accuracy of VAV system representation in whole-building energy simulation models. This study particularly aims at potential inadequacy of the EnergyPlus representation of an air-handling unit (with variable speed fans) connected with terminal VAV boxes (with reheat capability). The remaining of this section are structured as the following: First, we present a schematic of the VAV system we will analyze, and with this baseline established, we can readily identify physical behaviors neglected in the EnergyPlus model. Then in order to quantify the impact of such simplification, we construct a high fidelity model that captures missing physical behavior. The next step will be an effort of quantifying whether there is significant discrepancy between the low and high fidelity models and if the outcome of the high fidelity model (or the discrepancy between both models) could be mapped onto existing parameters within EnergyPlus and represented by their uncertainty. Eventually, we analyze how the results can be generalized for uncertainty analysis of future VAV systems.

4.4.2 Schematic of an Air-Handling Unit with Terminal VAV Boxes

We choose to study a variable air volume (VAV) flow system with economizer and a heating and cooling coil in the air handling unit, and there are terminal VAV boxes

with reheat coils at each supply branch. This is the common setup of a VAV system, but the control sequence of the same system setup could lead to quite different implementations. For instance, ASHRAE (2005) suggests two control sequences for the chilled water coil/hot water coil/economizer system type, and there are in total 19 sequences for VAV systems depending on system configurations. In this study, we implement the control sequence VAV 2A2-21232. The system schematic along with control system objects are depicted in Figure 4.18.

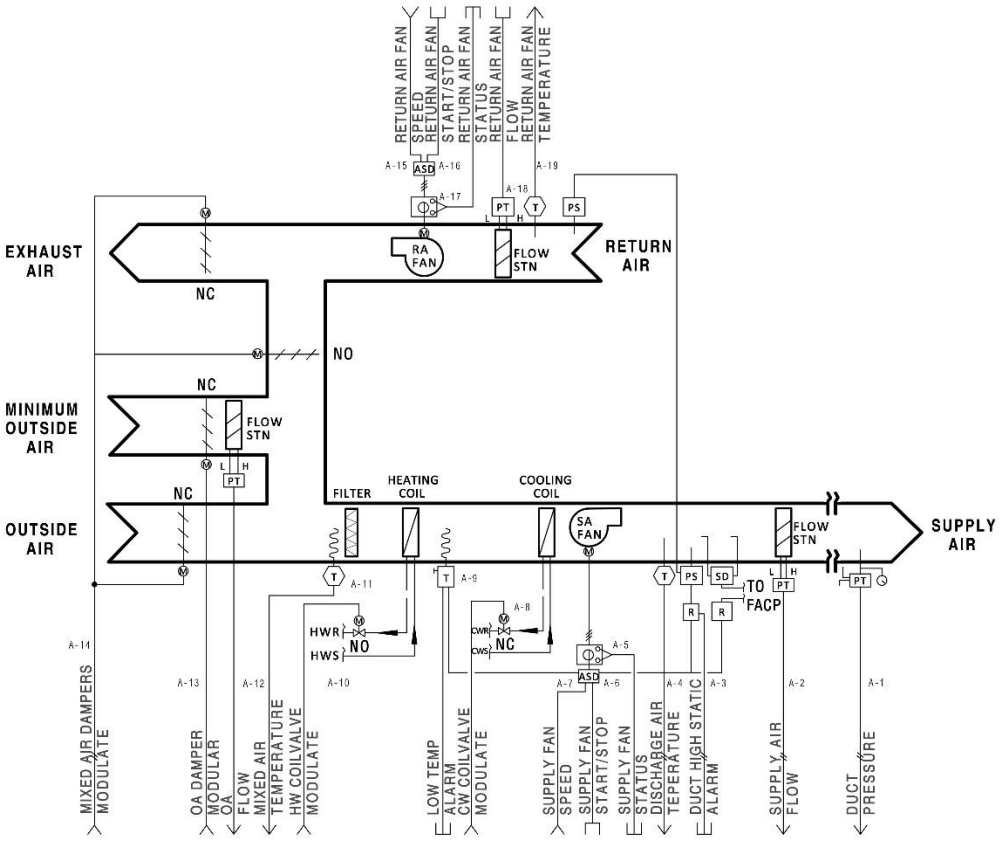


Figure 4.18 Schematic for control sequence VAV 2A2-21232 (Adapted from (ASHRAE, 2005))

The air handling unit supply fan speed modulates to maintain duct static pressure set-point, but the minimum speed is 30% of maximum to ensure sufficient fan motor cooling, while the return fan is regulated to track the flow rate of the supply fan, or

reduced by a fixed amount to account for local exhaust fans. The duct static pressure is reset based on terminal VAV box demands so that at least one of the VAV boxes is at 90% open. The heating coil valve, mixed air dampers and cooling coil valve modulate and together maintain the supply air temperature. Economizer free cooling is enabled when outside air dry bulb temperature is less than the return air's minus a deadband, otherwise the maximum outside air dampers shall close and the minimum outside air damper modulate to maintain the minimum outside air flow set-point. In each supply branch, the VAV damper is controlled to meet the room temperature set-point for cooling, or fully open during heating. The reheat coil is adjusted to track the room temperature set-point for heating.

4.4.3 Description of the EnergyPlus Model and Its Potential Inadequacy

The major simplification of an EnergyPlus (version 8.2) representation of the above system is that EnergyPlus does not perform pressure-based simulations of the air loop. Such simplification has several repercussions: first, all VAV boxes in EnergyPlus operate standalone, without interacting with other VAV boxes, while in reality, when one VAV damper opens or closes, it might impact the duct static pressure distribution as well as static pressure set-point and therefore the fan speed, such that all remaining VAV dampers have to adjust accordingly. In the actual VAV system, all airflows are determined by the duct static pressure distribution and the performance curves of the fans. Second, since the system pressure that the fan is facing is not known, there is no way of computing the fan speed, which is required for computing the electrical power accurately using the fan similarity laws. In fact, the implementation of variable speed fan within EnergyPlus is quite simple:

$$f_{flow} = \dot{m}/\dot{m}_{design},$$

$$f_{pl} = c_1 + c_2 \cdot f_{flow} + c_3 \cdot f_{flow}^2 + c_4 \cdot f_{flow}^3 + c_5 \cdot f_{flow}^4,$$

$$Q_{tot} = \frac{f_{pl} \cdot \dot{m}_{design} \cdot \Delta P}{\eta_{tot} \cdot \rho_{air}}, \quad (4.6)$$

where

Q_{tot} is the fan power in watts,

$\dot{m}$ is the air mass in kg/s,

$\dot{m_{design}}$ is the design air flow in kg/s,

ΔP is the fan design pressure rise in Pascals,

η_{tot} is the fan total efficiency,

f_{flow} is the flow fraction,

f_{pl} is the part load factor.

Analyzing input parameters for the fan model, we realize that $\dot{m_{design}}$ is the sum of zone design airflows from the EnergyPlus sizing calculations, and e_{tot} is information from the manufacture. ΔP is associated with uncertainty, since the system resistance that the fan faces at design airflow is not known precisely for the modeler. In addition, the polynomial curve that relates f_{pl} with f_{flow} is also a matter of choice for the energy modeler. Table 4.7 below, from Wray and Matson (2003), suggests default coefficient values, but the exact condition under which they are derived may not match the situation in reality. In the study, the one for the generic variable speed drive is chosen as a baseline for further comparisons, since this curve is also used by the EnergPlus default HVAC template for VAV systems. The third caveat is that there are no local controls such as proportional and proportional-integral controllers implemented in EnergyPlus. Instead, each branch or damper is operated in an idealized fashion, in the sense that the supply side is always able to deliver what is demanded subject to component capacity from sizing calculations, and that dampers are able to control airflow perfectly without any delay or fluctuations.

Table 4.7 Polynomial coefficients for fan performance curves

Fan control type	C1	C2	C3	C4	C5
Outlet damper	0.3507	0.3085	-0.5414	0.8720	0

Inlet vane	0.3707	0.9725	-0.3424	0	0
Variable speed drive (generic)	0.0015	0.00521	1.1086	-0.1164	0
Variable speed drive (Title 24)	0.1021	-0.1177	0.2647	0.7600	0

4.4.4 Description of the High Fidelity Model

We adapt a Modelica implementation of a VAV system from the Modelica Buildings library and apply to this study. Modelica is an equation-based object-oriented modeling language that specializes in describing and analyzing complex systems represented by differential, algebraic and discrete equations. The Buildings library (Wetter, Zuo, Nouidui, & Pang, 2014) is developed by LBNL to support rapid prototyping of innovative design and control of building systems, and has been widely received as of high fidelity. In the Modelica model of the VAV system, we capture actual control sequences, in the sense that the control input is always a physical quantify measured by a sensor, compared to an idealized thermal load customary in many building simulation programs including EnergyPlus. In addition, the control output is an actual control signal, such as the damper or valve position, contrary to flow rate in EnergyPlus. Therefore, the Modelica model can accurately represent air flowrate at a given VAV box as not only a function of its own damper position, but also of the damper characteristics, upper stream fan speed, fan curve, and the entire duct distribution network including other VAV box dampers. We highlight some more details of this model regarding the modeling of fans, duct resistance and damper resistance.

4.4.4.1 Fan Performance Curve

In this study, we implement a forward-curved fan from Carrier. Figure 4.19 (left) shows the fan performance curves at the maximum and minimum speeds, normalized by the design airflow rate on the X-axis, and the design static pressure on the Y-axis. Figure 4.19 (right) shows normalized fan total efficiency (with regard to maximum efficiency at

0.611) against normalized airflow rate. During the simulation, once the control output in terms of normalized fan speed r_N (0 to 1) has been determined by the controller, for instance, tracking the duct static pressure set-point, the new fan operating point is governed by the following equation that relates fan static pressure with volume flow rate:

$$\Delta P = r_N^2 \cdot s\left(\frac{\dot{V}}{r_N}, d\right) - \Delta P_\varepsilon ,$$

where

$\dot{V}$ is the volume flow rate in m^3/s,

s is a cubic hermite spline that interpolates input data points d on the performance curve with $r_N = 1$.

Similarly, the following equation computes fan efficiency based on fan speed and volume flow rate:

$$\eta = \eta_{max} \cdot s\left(\frac{\dot{V}}{r_N}, d\right),$$

where s is a cubic hermite spline that interpolates input data points d on the efficiency curve with $r_N = 1$, η_{max} is the maximum fan total efficiency.

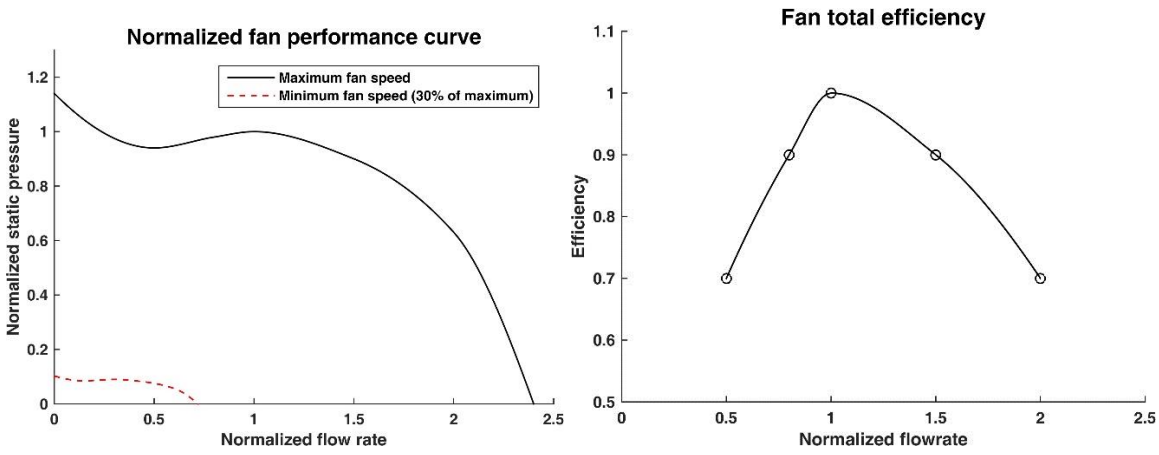


Figure 4.19 Fan characteristics in this study

When the fan pressure, flowrate and efficiency are all determined, Equation 4.6 is also used for computing fan power consumptions.

4.4.4.2 Duct and Damper Resistance Characteristics

The static pressure loss when the air stream flows through ducts can be described with a resistance model with a fixed flow coefficient, governed by the following equation:

$$\dot{m} = k\sqrt{\Delta P},$$

where k is a constant flow coefficient in $\sqrt{kg.m}$.

In contrast, the flow resistance of a damper varies with the damper position. In this study, we implement a damper model from ASHRAE 825-RP (Haves, Norford, DeSimone, & Mei, 1996) with its flow coefficient an exponential function of the opening angle. Damper flow resistance can be represented by:

$$k_d(y) = \exp(a + b(1 - y)), \tag{4.7}$$

where

k_d is the damper characteristics,

a and b are constants,

y is damper position, when $y = 0$ corresponds to a closed damper, while $y = 1$ means the damper is open.

The above equation only applies when damper position is within range bounded by (y_L, y_U). Outside this range, k_d is defined by a quadratic polynomial that matches the damper resistance at $y = y_L$ and $y = 0$, or $y = y_U$ and $y = 1$, respectively. The flow coefficient of the damper can be calculated as:

$$k(y) = \sqrt{\frac{2\rho}{k_d(y)}}\, A, \tag{4.8}$$

where A is the face area of the damper.

Since the resistance of a VAV box is a combination of a fixed-resistance component and a time varying one, we combine Equation 4.7 and 4.8 and derive the flow coefficient for a VAV box:

$$k_{eff} = \frac{k_{fix} \cdot k(y)}{\sqrt{k_{fix}^2 + k^2(y)}},$$

where k_{eff} is the effective flow coefficient for a VAV box.

4.4.5 Uncertainty Quantification of the EnergyPlus VAV Model

Now that we have implemented a Modelica representation of the VAV system, we are ready to investigate the potential inadequacy of the EnergyPlus VAV model. In order to enable a fair comparison between both HVAC system representations, we need to ensure consistency between the building thermal responses of both models. First of all, we take as a case study one floor from the new construction medium office building in Atlanta from the DOE reference building pool (Deru et al., 2011), which consists of a core zone and four perimeter zones (southern, eastern, northern and western zones). Second, in the EnergyPlus simulation, slight update of the HVAC system in the medium office building is conducted to ensure that the HVAC representation matches closely the VAV system in Figure 4.18. At the same time, we remove all HVAC related objects from the medium office building model and perform a co-simulation within Modelica, where the building thermal model interacts with a high fidelity HVAC implementation. Details of the co-simulation are summarized below.

4.4.5.1 Co-Simulation of the EnergyPlus Building with Modelica HVAC System

Preparation for a co-simulation of EnergyPlus building and Modelica HVAC system takes the following three steps: configuring the EnergyPlus model for co-simulation (Nouidui, 2014), exporting the EnergyPlus building as a Functional Mock-Up Unit (FMU), configuring the Modelica model for co-simulation. An FMU is a simulator or model that supports co-simulation according to Functional Mock-Up Interface (Pazold, Burhenne, Radon, Herkel, & Antretter, 2012), which is a tool-independent standard for model exchange and co-simulation of dynamic models.

4.4.5.1.1 Configuring the EnergyPlus Model

The role of the EnergyPlus model is to capture the thermal response of the room resulting from supply and return airflow by the Modelica HVAC system, and provide feedback to the Modelica model with room air temperatures. Therefore, in EnergyPlus, we create "OtherEquipment" objects associated with schedule values being the sensible and latent heat gain resulting from the supply air flowing into the room boundary and the return air flowing out. During the co-simulation, the Modelica model will dynamically change these schedules values and feed into EnergyPlus, while the latter will output room air temperatures to be read by Modelica controllers. A simplified configuration of EnergyPlus with schedules is shown in Figure 4.20.

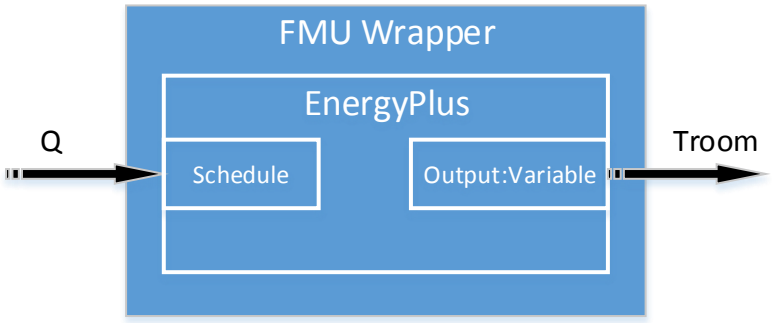


Figure 4.20 Configuration of EnergyPlus with schedules

4.4.5.1.2 Exporting the EnergyPlus Building as an FMU

For exporting the EnergyPlus model as an FMU, we adopt the utility "EnergyPlusToFMU", developed by LBNL (LBNL Simulation Research Group, 2016). The resulting FMU may contain the FMI model description file, the C sources of the FMU and additional FMU data. Now we have a system with two simulation programs: EnergyPlus being the slave program packaged as an FMU for co-simulation, and Modelica as the master program that imports the FMU. The detailed data exchange algorithm is described below.

Let N denote the total number of time steps and let k denote the current time step. We use f_1 and f_2 to denote the functions that compute the next values of the state variables x of the simulator 1 and 2, respectively.

Then program 1 computes the sequence

$$x_1(k+1) = f_1(x_1(k), x_2(k)),$$

and, similarly, program 2 computes the sequence

$$x_2(k+1) = f_2(x_2(k), x_1(k)).$$

At the end of each time step, each program sends its state variable values to and receives from the other program. Although each program uses its internal time integration algorithm, the data exchange between them is analogous to an explicit Euler integration. Therefore, in situations where the differential equation is solved with co-simulation, the algorithm is the following:

Step 0:

Initialize counter $k = 0$ and number of steps, N.

Set initial state $x_1(0) = x_{1,0}$ and $x_2(0) = x_{2,0}$. Set the time step $\Delta t = 1/N$.

Step 1:

Compute new states

$x_1(k+1) = x_1(k) + f_1(x_1(k), x_2(k))\Delta t$, and

$x_1(k+1) = x_1(k) + f_1(x_1(k), x_2(k))\Delta t$.

Replace k by $k+1$.

Step 2:

If $k = N$ stop, otherwise go to Step 1.

Note that the above algorithm allows no iteration between both simulation programs within one time step, so in essence the coupling scheme is based on loose coupling.

4.4.5.1.3 Configuring the Modelica Model

Recall that the EnergyPlus FMU contains only the building model without HVAC systems, so the Modelica HVAC model cannot export the amount of supply and return airflow to EnergyPlus. Instead, we need to compute within Modelica the amount of

sensible and latent heat flow into each room boundary by considering both supply and return air streams. In addition, since the medium office building does not perform a pressure-based airflow network simulation, the coupling scheme in our co-simulation ignores the pressure disturbance to the room by HVAC airflows as well as infiltration and inter-zone air flows. For now, we assume these simplifications do not significantly undermine the fidelity of our model.

In terms of implementation, we create a Modelica component that represents the return air stream from each room. Such a component takes as input the mass flowrate, air temperature, and relative humidity and outputs the sensible and latent enthalpy flowrate of the air streams with properties of the return air from each room, as depicted in Figure 4.21.

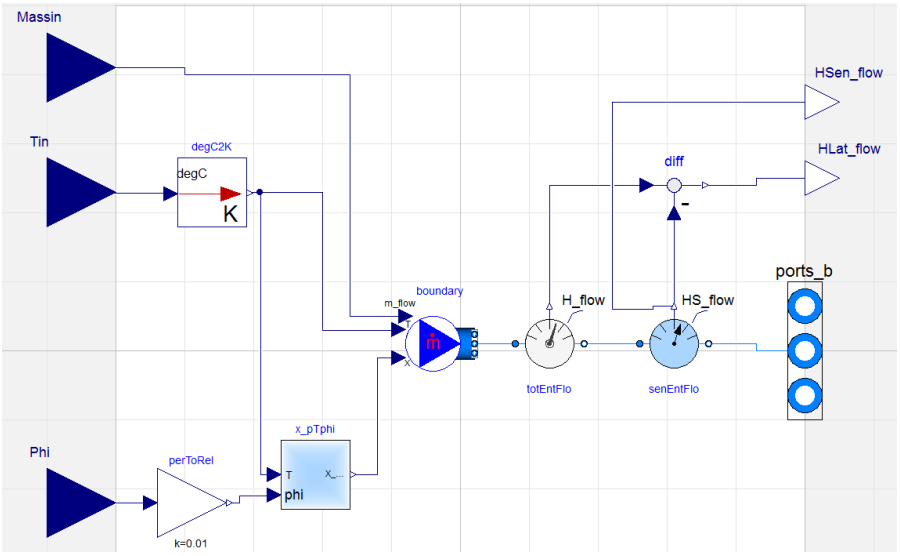


Figure 4.21 A specialized component in Modelica

The key part of this component is the "Buildings.Fluid.Sources.MassFlowSource_T" from the Buildings library, which produces a prescribed mass flow with prescribed temperature and mass fraction. As such, we can compute the difference in sensible and latent heat flowrate between the supply and return air streams for each room. The overall Modelica model is shown in Figure 4.22. Since the Modelica model does not have auto-sizing features, we take design supply airflows for each zone from EnergyPlus sizing

results and input into Modelica. In return, we take the fan operating pressure that corresponds to the HVAC ductwork and use it for EnergyPlus fan consumption calculation (ΔP) in Equation 4.6.

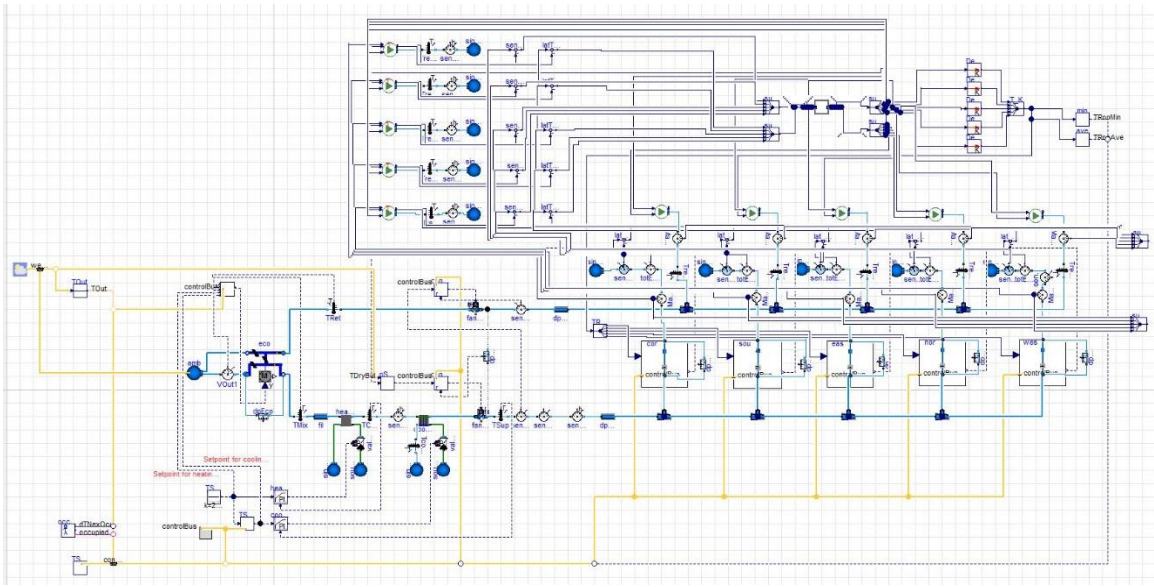


Figure 4.22 Overall Modelica model for the VAV system

4.4.5.2 Uncertainty Quantification

Before uncertainty quantification of the EnergyPlus model, we demonstrate that the high fidelity co-simulation is able to capture physical behaviors neglected in the EnergyPlus model.

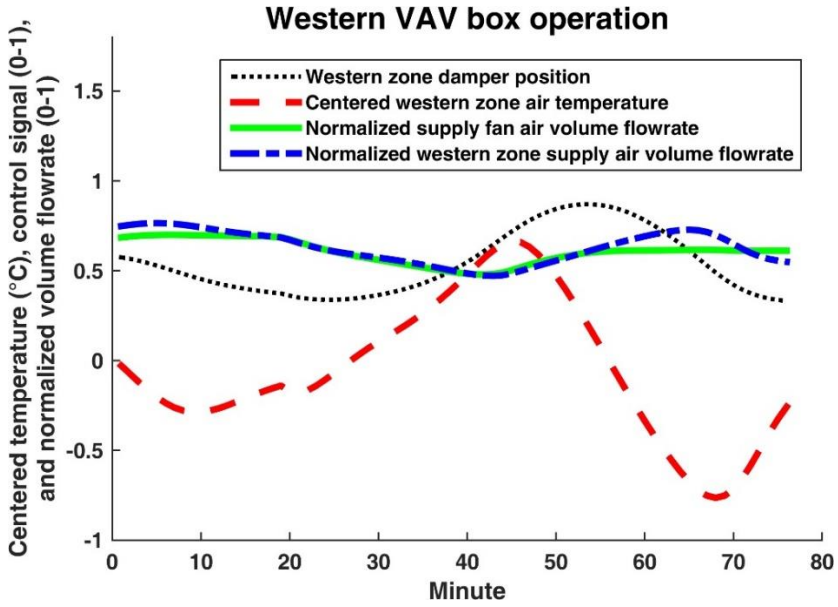


Figure 4.23 Western VAV box operation

In Figure 4.23 that depicts a sunny afternoon on July 5th, air volume flowrates are normalized against design airflow rates, rooms temperatures are centered around the cooling set-point at 24 ℃, and zone damper position is always between zero and one. If we zoom in and observe interactions between the western zone VAV box with the rest of the system, we find that from around Minute 28, the western room suffers from overheating, so the PI controller of the VAV box instructs the damper to open up, asking for more supply air into the zone. However, at this time of the afternoon, the ambient air is sufficiently cool such that all other zones demand less supply air, so the supply fan reduces its speed and overall supply air flowrate keeps decreasing. Consequently, the western zone is not supplied with sufficient amount of supply air despite of the increasing demand by the VAV box, and the room air temperature overshoots 0.7 ℃ above the set-point until Minute 45. Afterwards, the damper opening angle continues to increase since the room air temperature is still above the set-point. Such interactions cannot be captured by the EnergyPlus model since it assumes each VAV branch immediately receives the desired amount of supply air as long as the whole system capacity is not exceeded.

Figure 4.24 below (volume flowrates are normalized with design supply volume flowrate) showcases an issue with the current control sequence, when the supply fan speed is sufficient low resulting from the throttling of all VAV dampers on July 28 with mild weather, the amount of outdoor air supplied may drop below prescribed system minimum outdoor air requirement. This observation is in line with a statement from ASHRAE (2005), describing the performance of the particular control sequence: "At low fan speeds the mixed air plenum may not be negative enough to draw in sufficient outside air volume".

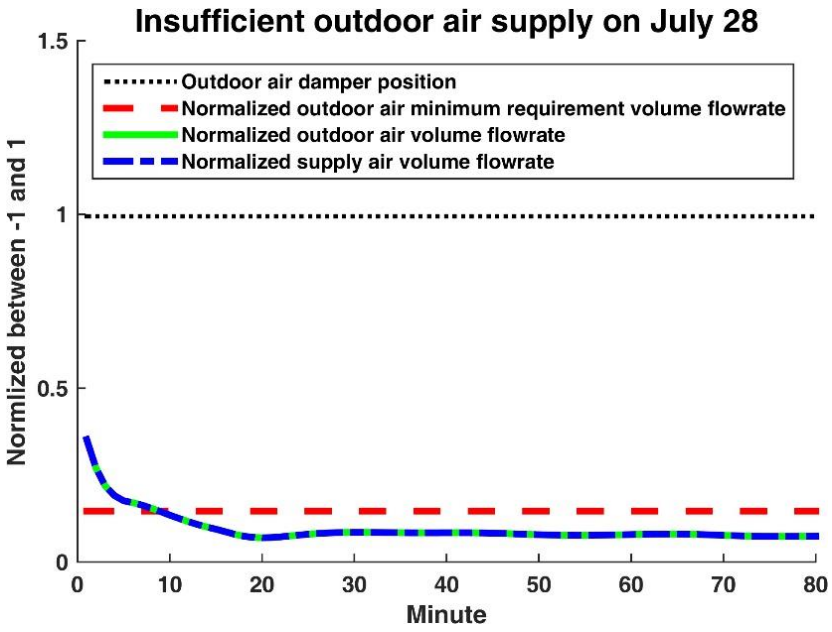


Figure 4.24 Drawback of the current control sequence

Now we compare the simulation results of both platforms for the whole month of July in Atlanta in Figure 4.25.

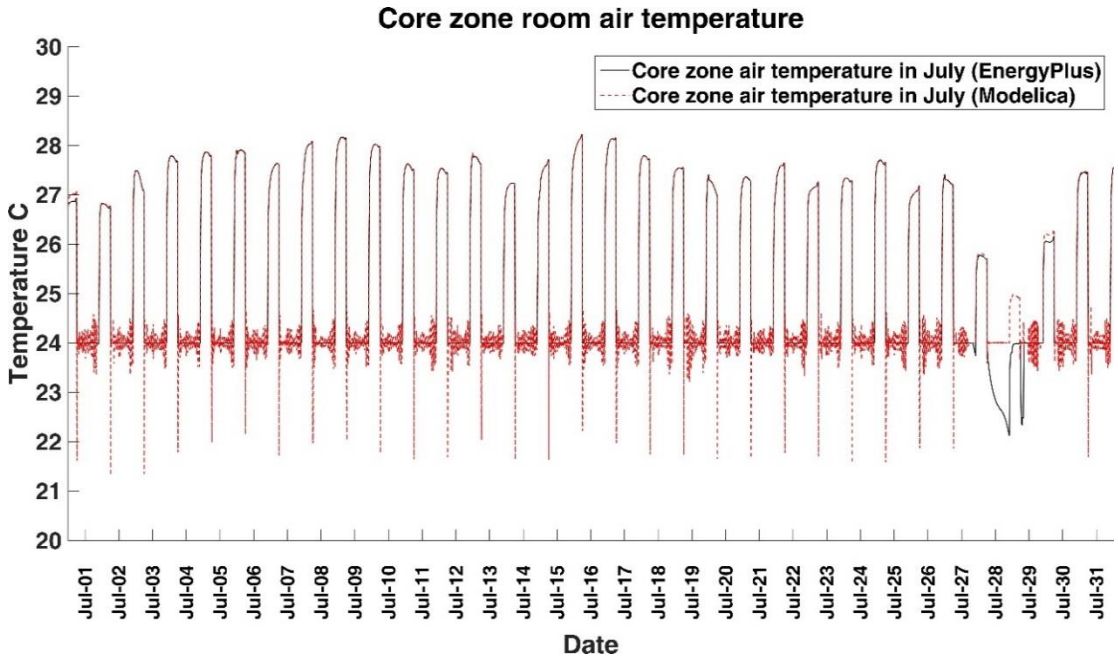


Figure 4.25 Comparison of core zone room air temperature

On all days but July 28, the temperature profile predictions for the core zone from both EnergyPlus and Modelica models follow almost the same trend, except that the temperatures from the Modelica model fluctuate around the set-point of 24 ℃ resulting from the imperfect control. Normalized mean bias error ($NMBE = \frac{\sum_1^n (y_i - \widehat{y_i})}{(n-p)\cdot \bar{y}}$) of the EnergyPlus model is only 0.14% if we regard Modelica results as a high-fidelity baseline.

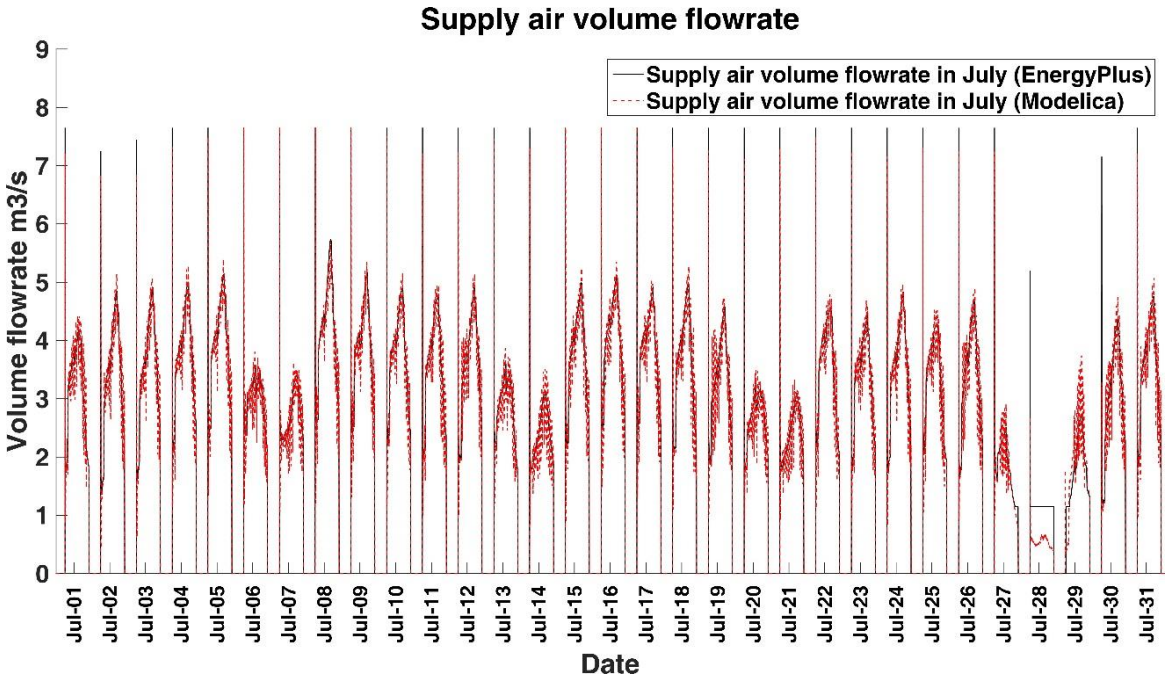


Figure 4.26 Comparison of supply air volume flowrate

In terms of supply air flowrates in Figure 4.26, both platforms capture the day-to-day variability well, and the $NMBE$ of EnergyPlus is only 1.3%. From the previous realization that EnergyPlus system representation has neglected lots of physical behaviors of the actual system, one might find it puzzling how the result of EnergyPlus comes out so close to the high fidelity model. However, this is expected if we realize that both HVAC systems deliver supply air at the same state into the same building, and the room temperature responses show little discrepancy, so the predictions for supply air flowrates from both platforms should match well. If we zoom in and observe the temperature and supply air flowrate comparisons in Figure 4.27, we realize the within-day fluctuations shown in the high fidelity model are the results of VAV box interactions and imperfect control, which do not play a role in aggregated metrics such as $NMBE$. Meanwhile, since our variable of interest is still energy, we do not expect EnergyPlus to capture these within-day fine details.

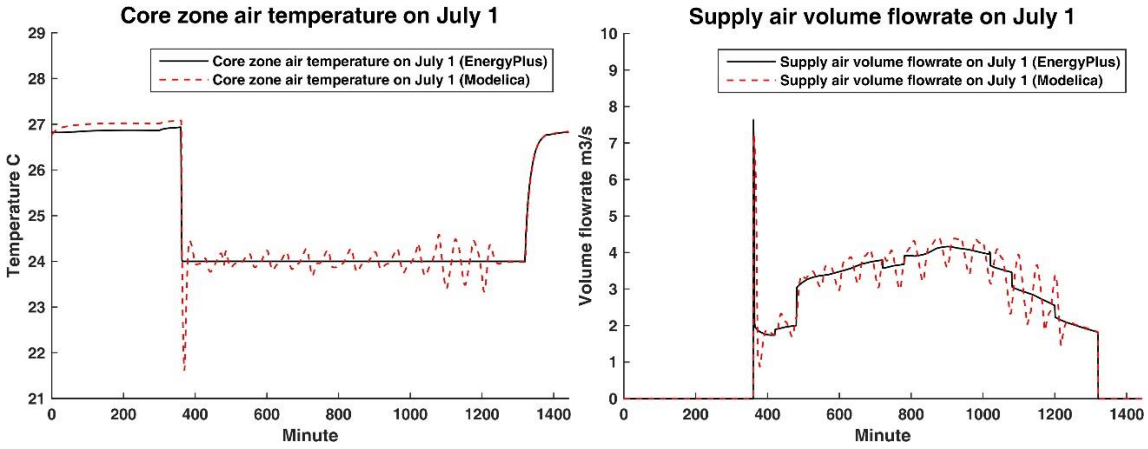


Figure 4.27 Detailed temperature and supply air flowrate profile

When it comes to the comparison of fan power in Figure 4.28, EnergyPlus results show significant discrepancy with the high fidelity model results. Causes of such discrepancy are twofold: first, EnergyPlus simulates neither the fan speed nor the efficiency fluctuations with the varying load; second, EnergyPlus adopts default polynomial curve that relates f_{pl} with f_{flow}. In this case, EnergyPlus has a $NMBE$ of 41.5%.

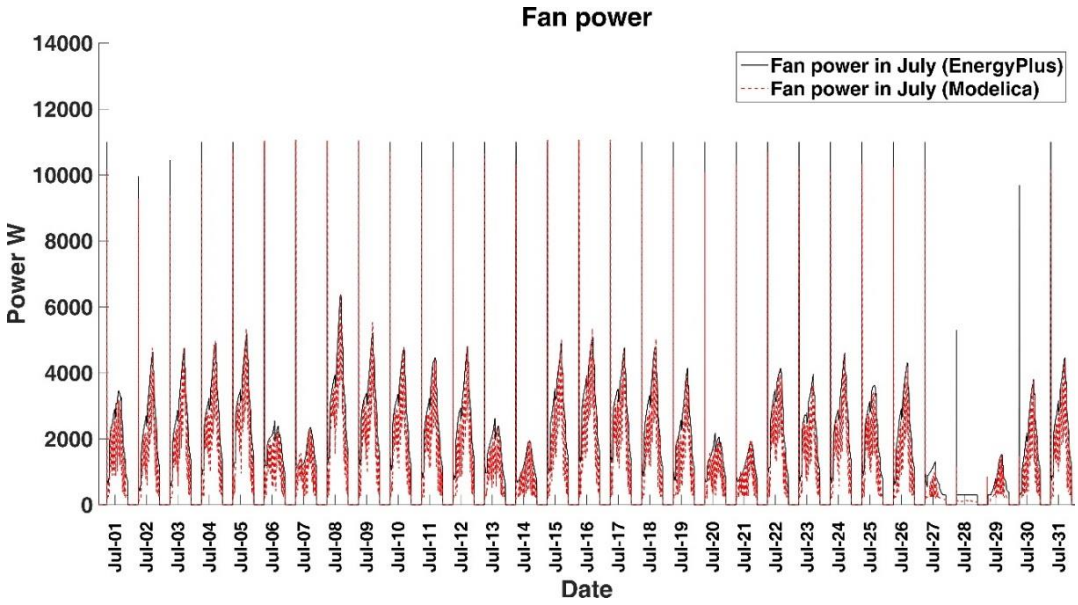


Figure 4.28 Comparison of fan power

Now that we have realized the inadequacy of EnergyPlus for predicting fan power accurately for the type of VAV system we are studying. In order to quantify this as a

model form uncertainty, we can either map the high fidelity model to existing parameters from EnergyPlus or quantify a statistical model for the discrepancy term that relates both platforms. In this situation, since EnergyPlus predicts supply air flowrate quite well, and its built-in polynomial curve has several exposed parameters to work with, we choose the former approach and map the high fidelity model to existing parameters from EnergyPlus. To do so, we perform a least square polynomial curve fitting to normalized fan power (relative to fan power at design condition) against normalized volume flowrate from the Modelica model. Results are shown in Figure 4.29, and the fitting R^2 is 0.974.

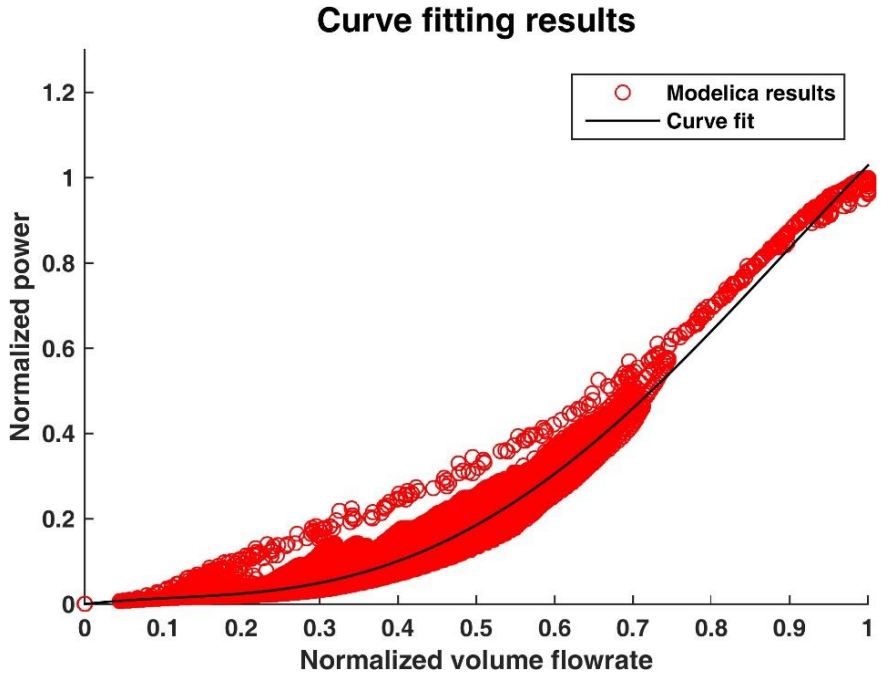


Figure 4.29 Curve fitting results

The fitted polynomial coefficients are (0, 0.2169, -1.186, 3.976, -1.978). As mentioned previously, this curve fit subsumes the combined effect of fan model simplification and imperfect control. If we rerun the EnergyPlus model with the updated coefficients, the results are shown in figure 4.30 with a much lower $NMBE$ of -1.37%.

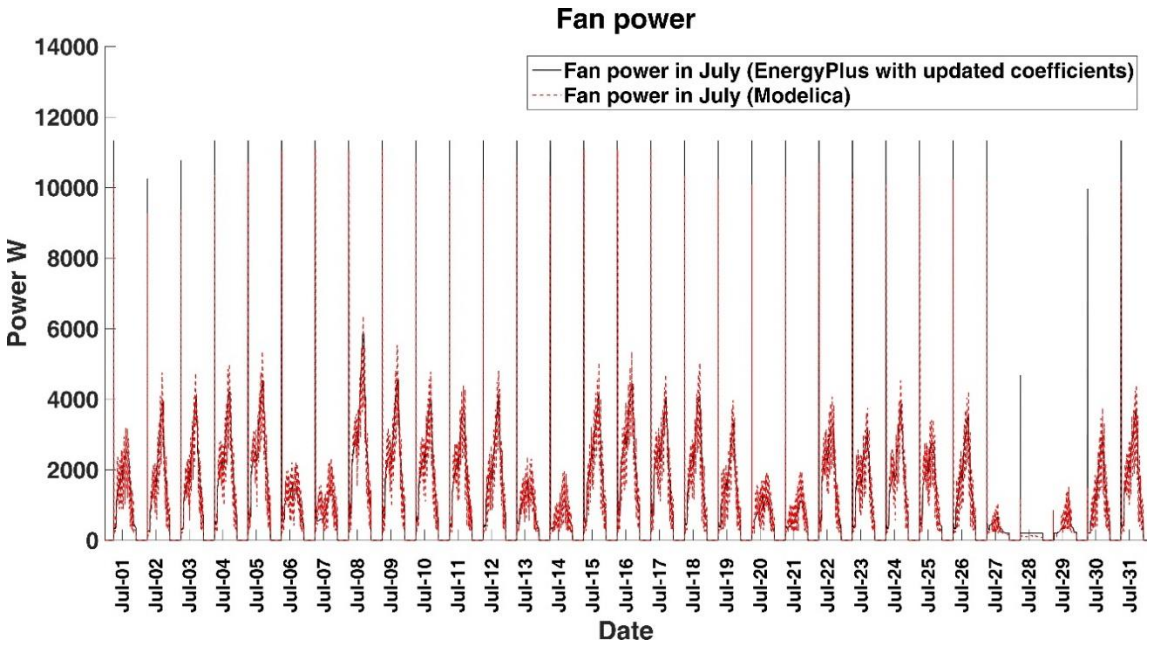


Figure 4.30 Comparison of fan power with updated coefficients

From Figure 4.29, we observe that the residuals (the discrepancy between actual and fitted values) from the statistical curve fit seem to show structural patterns, for instance, variance of the residuals are large at lower normalized flowrate than that at higher flowrate, which suggests issues with the statistical goodness of fit. The underlying reason for the structural discrepancy is that at lower fan speeds, the VAV damper open angles adjust frequently to varying loads at each zone, such that a specific normalized volume flowrate might correspond to a large variability of fan speeds and system resistance. On the contrary, at higher fan speeds, VAV dampers are mostly open such that the system resistance is closer to design conditions. Such an effect is evident in Figure 4.31, where we compare the predicted fan operating points from Modelica as well as from EnergyPlus with the updated coefficients. Since EnergyPlus does not perform pressure-based simulation, we simply divide normalized power by normalized volume airflow for effective normalized pressure rise. Eventually, we choose to neglect the discrepancy observed at such fine details, since the focus of this polynomial fit is to capture aggregated energy consumption of fans, where $NMBE$ as an adequate measure for the overall performance of the statistical model approves the fitted coefficients.

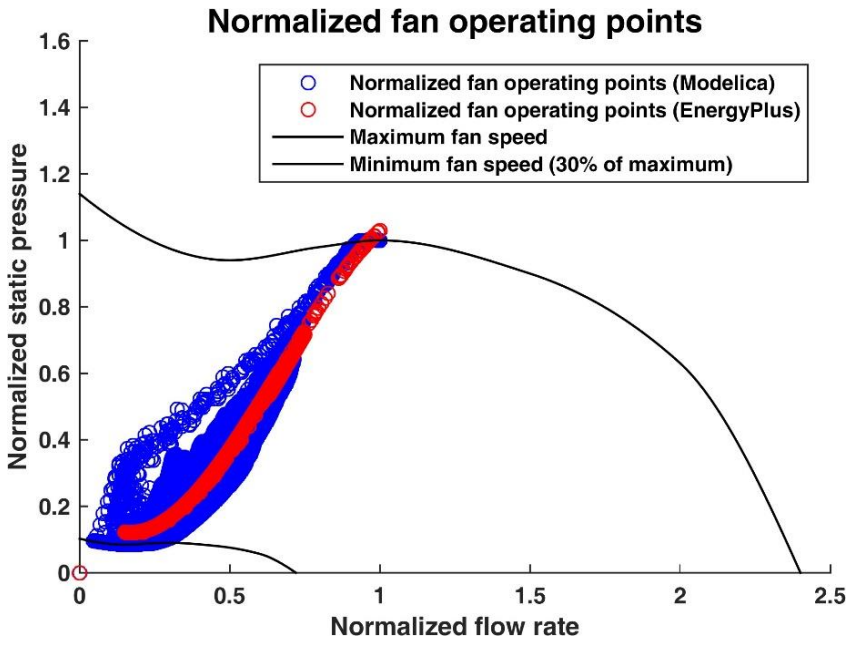


Figure 4.31 Predicted fan operating points by both models

4.4.6 Result Generalization for Future Analysis of VAV Systems

So far, we have attempted the quantification of model form uncertainty associated with the EnergyPlus model of VAV systems. We know the resultant polynomial curve at least applies to the particular medium office building in Atlanta and the particular ductwork network in the Modelica model. However, in order to reach conclusions that are more general, we conduct parametric variations to the current case and analyze results.

The first variation is to change the building model: will the curve fit still apply if the building thermal demand changes? We update the building thermal model with higher internal loads, subject to hotter weather at Phoenix. Since the thermal demand of the building changes, we perform another sizing calculation in EnergyPlus and update design fan pressure rise according to a Modelica HVAC simulation. However, the normalized fan performance curve and system resistance characteristics remain the same. Results for the fan power comparison are shown in Figure 4.32, which present an excellent match. The $NMBE$ is only -0.89%.

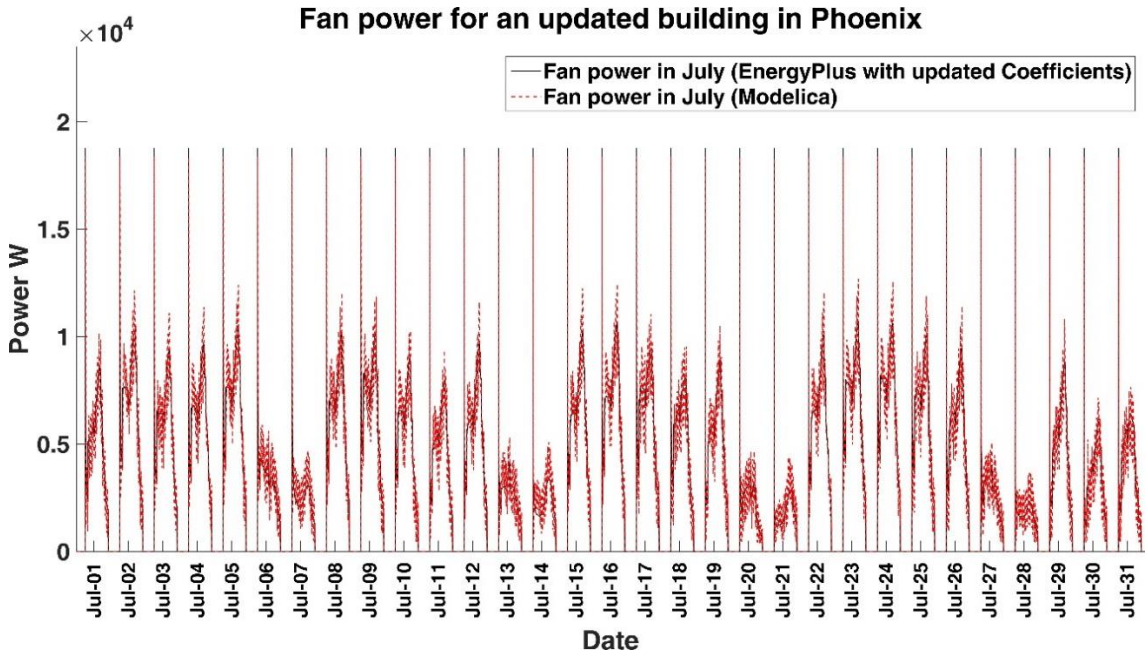


Figure 4.32 Comparison of fan power for an updated building in Phoenix

The second variation is to change the system resistance characteristics: will the curve fit still apply if the building ductwork changes? To study this, we shrink the ductwork and damper flow coefficient k to be 80% of the original assumptions, so effectively the fan is now facing a ductwork network with higher resistance resulting from higher friction loss and smaller face areas at the VAV dampers. Again, the airflow sizing calculation and fan sizing simulations are conducted to make sure the fan delivers design airflow rate at design pressure rise in Atlanta weather. Together with previous simulation runs, only the normalized fan performance curve always remains intact. Results for the fan power comparison are shown in Figure 4.33, again with a great match. The $NMBE$ is -1.72%.

The lesson from these two additional parametric studies is that the previously derived polynomial fit seems to apply to buildings served by VAV systems associated with similar normalized fan performance curves and control sequence to the one we study in this section.

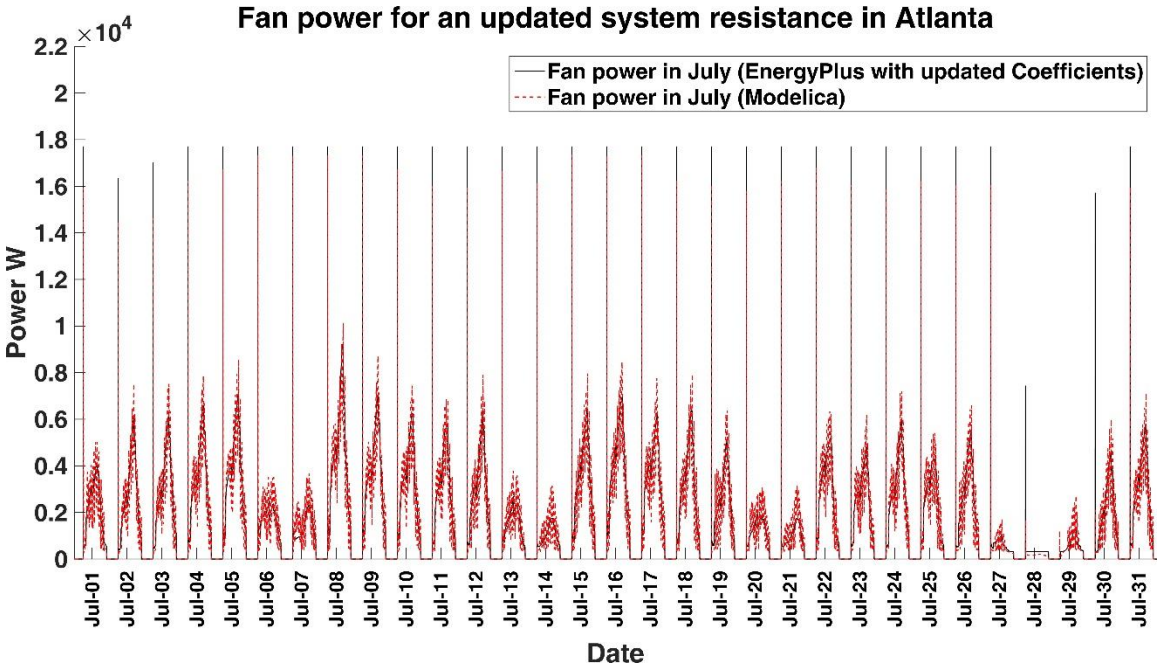


Figure 4.33 Comparison of fan power for updated system resistance in Atlanta

4.4.7 Realization Uncertainty of Duct System Resistance

Previously, we always perform pre-processing simulation in Modelica and size the pressure rise of fans to ensure they deliver design airflows when all VAV dampers are wide open. This assumption is reasonable since it is common practice for the building owner to hire a balancing contractor to adjust fan speeds according to actual system resistance after fans get installed onsite. Figure 4.34 illustrates such a situation.

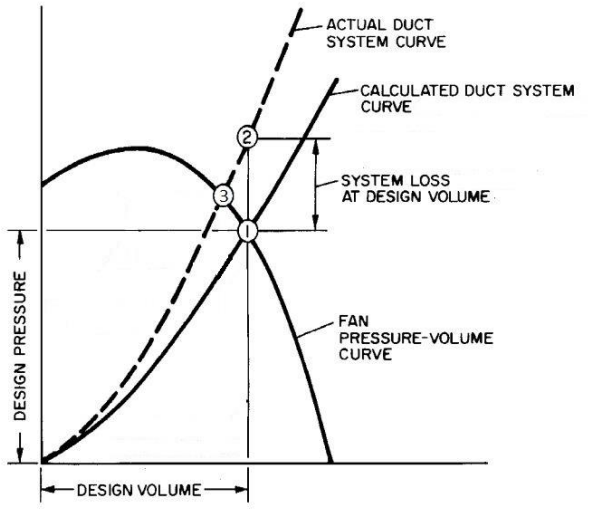


Figure 4.34 Predicted versus actual system pressure loss (adapted from (ASHRAE, 2001))

Due to time limitations, duct resistance and component loss can be largely overlooked during the design phase of the HVAC system. In practice, it is common for mechanical engineers to apply safety factors to account for the discrepancy between calculated and actual duct system resistance as depicted in Figure 4.34. According to the design duct system curve, the fan operates at Point 1 corresponding to the highest total fan efficiency, whereas in reality, realized ductwork may differ significantly from the design calculation, such that the fan actually operates at Point 3. In order to ensure the fan supplies sufficient airflow at design conditions, the balancing contractor may increase fan speeds such that the actual operating point is at Point 2. By operating at higher fan speeds, the fan not only operates astray from the maximum efficiency point but also needs to work with a larger motor, otherwise the design airflow cannot be met, which may cause thermal comfort issues. Therefore, it is imperative that mechanical engineers accurately account for system pressure loss, especially with fan system power limits now mandated by ASHRAE Standard 90.1-2010 (ASHRAE, 2010).

The remainder of this section aims at a preliminary attempt at quantifying the discrepancy between duct system resistance calculated at design stage and the actual in

situ resistance. This falls under realization uncertainty in the uncertainty taxonomy from Chapter 3. In the following texts, we present literature review, methodology and results.

4.4.7.1 Literature Review

The theory and data behind duct and fitting loss have been established reasonably well. The following is a brief summary of literature relevant to our study. Friction loss when air flows through ductworks can be calculated with the Darcy equation:

$$\Delta p_f = \frac{1000fL}{D_h} \cdot \frac{\rho V^2}{2},$$

where

Δp_f is friction loss in terms of total pressure in Pa,

f is friction factor, which depends on duct roughness factor and flow Reynolds number,

L is duct length in m,

D_h is hydraulic diameter in mm,

V is velocity in m/s,

ρ is density in kg/m^3.

For static pressure loss at local fittings, ASHRAE publishes a database (ASHRAE, 2011) with more than 220 round, flat oval and rectangular fittings. The local loss coefficient C is used for describing the ratio of total pressure loss to velocity pressure at the referenced cross section. Therefore, the total pressure loss in a duct section is calculated with the following equation:

$$\Delta p = (\frac{1000fL}{D_h} + \sum C) \cdot \frac{\rho V^2}{2}.$$

There are several complications, however, in the estimate of in situ duct system resistance. First, compressed flexible ducts are sometimes used for final connections of diffusers and terminal boxes. Contractors, instead of cutting the necessary length of duct needed, they may compress it. Abushakra, Walker, and Sherman (2004) propose pressure drop correction factors for compressed ducts based on their compression ratio, but

flexible duct elbows should be avoided. Second, there is usually the issue of duct leakage within ductwork systems in commercial buildings. Wray, Diamond, and Sherman (2005) show a wide range of duct leakage rates (L/(s.m^2 surface area) at 250 Pa) from 0.3 to 41 by conducting in situ measurements, whereas the ASHRAE (ASHRAE, 2001) recommended leakage rate for unsealed ducts is only 2.5. Wray and Matson (2003) show, by conducting simulations, that the increase in annual fan energy is estimated to be 40 to 50% for a system with a total leakage of 19% at design conditions compared to a tight system with 5% leakage. This study, nevertheless chooses to neglect duct leakage for lack of information, and the fact that sealed ducts for instance, at leakage class 4 (ASHRAE, 2001), only leak about 1% airflow at 250 Pa static pressure. Third, design calculations sometimes make no allowance for system effects, which is a term that describes certain inlet and outlet conditions that adversely affect fan performance. System effect results from the difference in how the fan was tested compared to how it is installed. ASHRAE fittings database provides loss coefficients that described these effects as an approximation, as fans of different types or by different manufactures interact with the ductwork in various ways. Fourth, fitting local loss coefficients are tested standalone, whereas in reality, consecutive fittings can be installed closely coupled (less than six hydraulic diameters apart). Under these situations, the flow patterns differ from test conditions such that the combined effect of these coupled fittings is not known precisely. Fifth, another contributor of the difference between theoretical and actual system resistance is the quality of installation or workmanship. Blue prints and specifications are sometimes not in sufficient details, which leave room for interpretation by contractors when they install ductworks onsite. In addition, ductworks may have to be depressed to avoid obstructions, such as water pipes and beams.

4.4.7.2 Methodology

With the goal of quantifying the discrepancy between theoretically calculated and actual ductwork resistance, we should ideally conduct a study by comparing engineering calculations according to previous theory and in situ measurements of actual fan performance at design air flowrate. However, we realize that the research community possesses only sporadic data sources of data on this subject. Therefore, we rely on our own data and for now only analyze the characteristics of one building as an illustration of the process. We study actual ductworks in a research-teach mixed used building on the Georgia Tech campus. The building has three floors, with a floor area of about 2985 m^2. We on one hand study HVAC plan drawings for the supply air ductwork (for instance Figure 4.35), and on the other hand observe actual duct realizations inside the building. We then re-create models of both ductwork system in SketchUp (Figure 4.36) to facilitate quantification. Realizing the discrepancy between the design and actual ductwork, two distinct theoretical calculations will be performed that represent truthfully the resistance characteristics of each scenario.

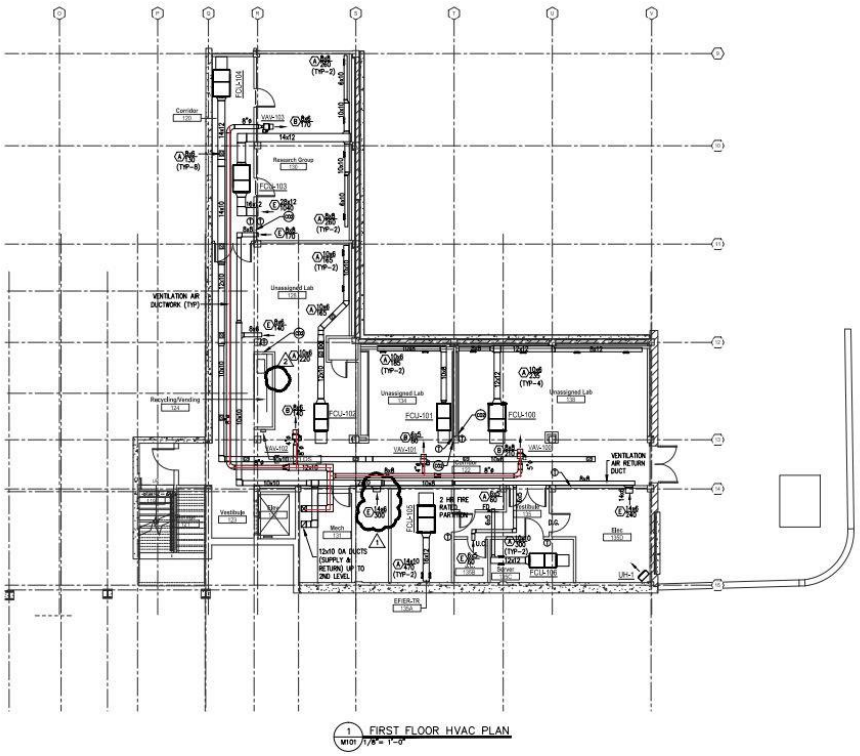


Figure 4.35 Supply air ductwork (highlighted) on the first floor

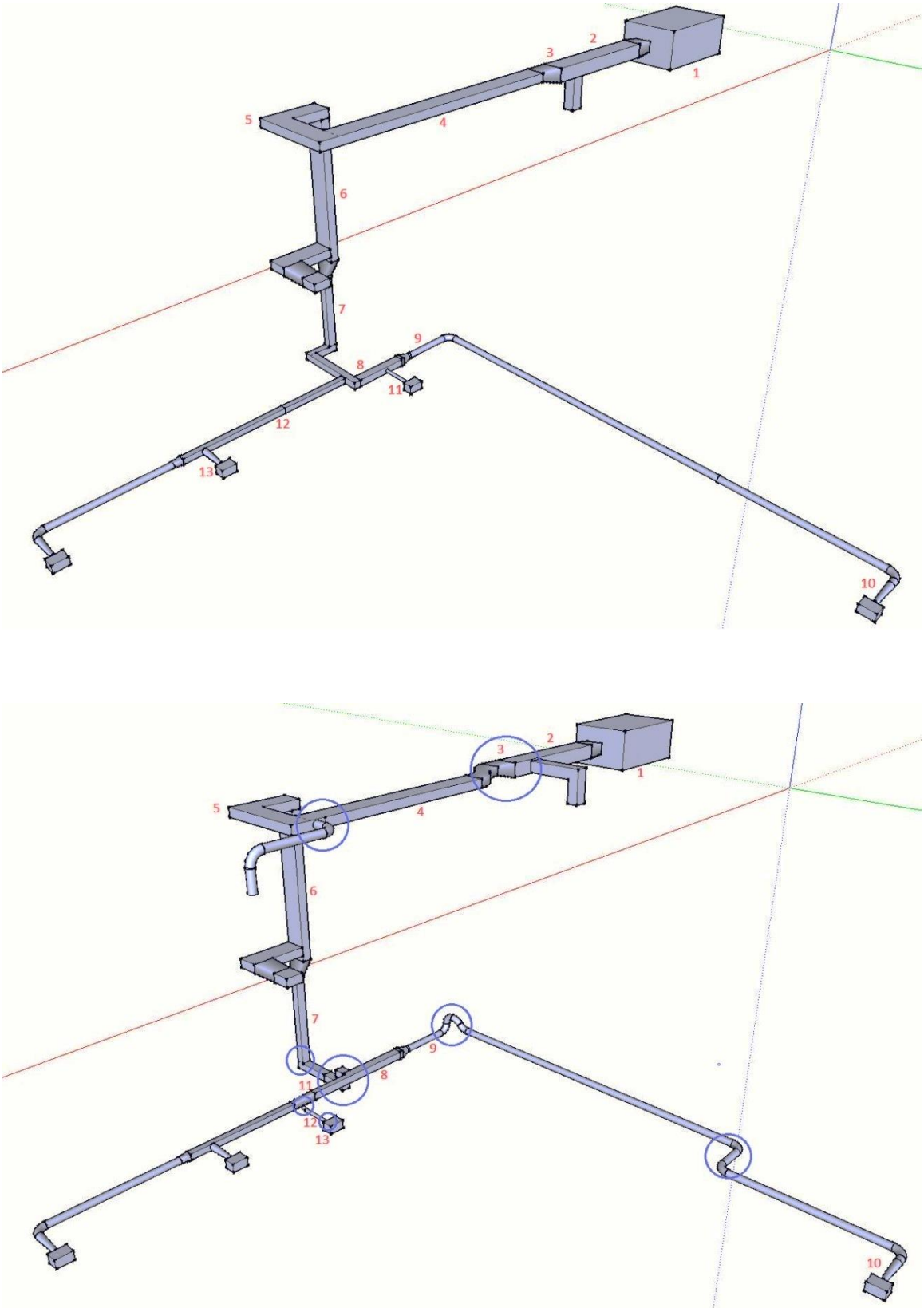


Figure 4.36 Design (upper) and actual (lower) duct layout

In Figure 4.36, the ductwork is divided into several sections according to the guideline from ASHRAE handbook (ASHRAE, 2001). In the actual ductwork layout, significant change from the design intention has been highlighted with circles. As expected, the ductwork from design idealization is much cleaner with less local loss compared to the actual realization. For instance, at Section 3 in the actual ductwork, the outlet main branch has to go through an irregular transition, probably because of requirement of service clearance space by the AHU. At the end of Section 4, the contractor decides to add another branch that serves the third floor underneath. At Section 7, the supply air duct for Floor 1 has to go through a z-shaped elbow and then a bullhead diverging tee, resulting from mechanical room space limitation. At Section 12, an additional elbow is added where the local air velocity pressure is high. At Section 13, compressed ducts are used for connecting the rigid section with the VAV box (Figure 4.37). Within Section 9, a few elbows are additional to the design scenario, for avoiding obstructions such as beams. Table A.1 and A.2 in the appendix detail the total pressure loss calculations by sections as well as loss coefficient summary by sections for the actual ductwork. The same is summarized in Table A.3 and A.4 for the design ductwork. Table 4.12 summaries the pressure loss of the critical paths for both scenarios. Results show that the external system resistance that the fan faces at design airflow rate (excluding internal resistance such as filter and cooling/heating coils) is 269 Pa in the design ductwork and 336 Pa in the actual one. Such discrepancy translates to 24.9% additional resistance not captured by a detailed ductwork resistance calculation at the design stage, resulting from not acknowledging onsite workmanship. If the fan is selected based on resistance according to the design situation, the fan either has to be adjusted to run at higher speeds than design, or will fail to deliver design airflow rate whenever needed. The range uniformly between 0% and 24.9% is identified as a preliminary estimate of the realization uncertainty associated with ductwork resistance, resulting from workmanship onsite. The lower bound, of course corresponds to a situation, where the contractor

installs the duct follows the blue prints and specifications "religiously". The limitation of the preliminary uncertainty quantification is that it fails to acknowledge the uncertainty with estimating the system effects and coupling of fittings, and it is only based on one case study building.

Figure 4.37 Additional elbow at an end branch and compressed ducts connecting the VAV box

4.4.8 Conclusions

We conclude this section about VAV systems by making the following recommendations for an energy modeler at the design stage when the fan selection has not been finalized yet. By adding an additional realization uncertainty factor between 0 and 24.9% to engineering calculations of the duct system resistance, and inserting polynomial coefficients derived from this study, the resultant uncertainty range for the fan power derived from the UA will contain the actual fan power. With recognition of both realization uncertainty and model form uncertainty, we have quantified the fan

power uncertainty in VAV systems. This result applies to forward-curved fans in VAV system similar to the ones in our study in terms of normalized performance curve and control sequence.

Table 4.8 Pressure loss of the critical paths for both actual and design Situations

Duct Section	Pressure Loss Actual Situation, Pa	Pressure Loss Design Situation, Pa
2	6.02	6.02
3	7.14	7.14
4	17.07	10.45
5	87.6	87.6
6	8.9	8.9
7	64.92	45.12
11	12.45	
12	69.47	14.56
13	62.87	89.64
Total	336	269

CHAPTER 5

SENSITIVITY ANALYSIS

Sensitivity analysis has gradually become an intrinsic part of uncertainty analysis for the identification of key input parameters affecting the prediction of building performance. In addition, since this dissertation puts a lot of emphasis on model form uncertainty, mixing their sensitivity into the picture allows for an inspection of the simulation engine for modules whose discrepancies rank high and therefore are most urgent to improve.

Literature on the topic of sensitivity analysis has not been scarce. de Wilde and Tian (2010) study the sensitivity of computational results to identify key design parameters, with rand regression and multivariate adaptive regression splines (MARS). Eisenhower, O'Neill, Fonoberov, and Mezic (2012) claim to extend traditional sensitivity analysis by studying the influence of about 1,000 parameters, using variance-based methods. Mara and Tarantola (2008) apply variance-based methods to a test-cell thermal model and show the hourly first-order and total sensitivity indices of the test cell's components on indoor air temperature. Mechri, Capozzoli, and Corrado (2010) also use variance-based methods to identify the design variables that have the most impact on the variation of the building energy performance for a typical office building. Pappenberger, Iorgulescu, and Beven (2006) propose a global sensitivity analysis method based on multiple regression trees (random forests). Ruiz, Bertagnolio, and Lemort (2012) also identify the most influential parameters affecting the final energy consumption in office buildings with variance-based methods. Spitz et al. (2012) apply successively local sensitivity analysis, correlation analysis, uncertainty analysis, and global sensitivity analysis to an experimental platform in France. Sun, Gu, et al. (2014) conduct parameter screening for removing insignificant parameters with lasso regression and then apply a

variance-based method for computing sensitivity measures for remaining ones. Tian (2013) gives a comprehensive overview of the current state-of-the-art methods.

Literature has not seen a rigorous methodology that ranks the sensitivity of input parameters along with model forms altogether. The need to do so is justified by the argument that spending effort in model improving (i.e. substitute low-fidelity modules with high-fidelity ones) and performing additional measurements for sensitive parameters are both means to the same end, which is to improve predictions, so their relative prioritization should be considered together. Furthermore, current methods only allow for the sensitivity analysis of individual parameters, but sometimes it is also interesting to see the overall sensitivity of a group of parameters, such as all parameters that influence the prediction of convective heat transfer coefficients. Therefore, we choose to apply group lasso with discrete categorical variables and sliced Latin Hypercube sampling.

5.1 Recap of Variance-Based Methods

Since variance-based methods are the most popular methods for sensitivity analysis, we first draw some background information from Saltelli, Tarantola, Campolongo, and Ratto (2004) for readers' information. The widely accepted variance-based methods yield robust and accurate global sensitivity measures without relying on any assumption on the nature of the input and output relations.

Assume that all the uncertainty parameters X are free to vary over the entire range of uncertainty, then the overall uncertainty of the outcome of interest $y = f(X)$ is measured by it unconditional variance $V(Y)$. The goal of a sensitivity analysis is often to rank the uncertainty parameters according to the amount of output variance that is removed when we learn the true value of a given input parameter X_i. Parameters could then be ranked with $V(Y|X_i = x_i^*)$, the variance obtained by fixing X_i to its true value x_i^*. Note that the $V(.)$ operator above is taken over all other parameters but X_i. We can normalize the measure by the unconditional variance $V(Y)$, which leads to $\frac{V(Y|X_i=x_i^*)}{V(Y)}$.

However, we soon come to the realization that the true value x_i^* of X_i is unknown, so it is natural to compute an average value of the above measure over all possible values of X_i. The parameter with the smallest $E(V(Y|X_i))$ should be ranked as the most influential to the outcome. In a rich notation, $E_{X_i}(V_{X_{-i}}(Y|X_i))$ denotes that the mean of the conditional variance, where X_{-i} stands for the vector of input parameters but X_i. Given that $V(Y)$ is a constant, and can be decomposed into $V(E(Y|X_i))$ and $V(E(Y|X_i))$ (law of total variance), searching for the smallest $E(V(Y|X_i))$ is equivalent to maximizing $V_{X_i}(E_{X_{-i}}(Y|X_i))$, which is the variance between the conditional means. As such, we define two sensitivity indexes as follows:

- First-order effect: $S_{X_i} = \frac{V(E(Y|X_i))}{V(Y)}$, where $V(E(Y|X_i))$ is expected reduction in the output variance that one would get if X_i could be known or fixed, or top marginal variance.
- Total effect: $ST_{X_i} = \frac{E(V(Y|X_{-i})}{V(Y)}$, where $E(V(Y|X_{-i})$ is expected residual output variance that would end up with if all factors but X_i could be known or fixed, or bottom marginal variance.

5.1.1 Features of First Order Effects

S_{X_i} is a good model-free sensitivity measure as it always gives the expected reduction in the variance of the output if one could fix an individual parameter. A nice property of them is that the sum of $S_X s$ is one, if the model is additive. A model $Y = f(X_1, X_2, \dots, X_k)$ is additive if the function f can be decomposed as a sum of k functions, and each f_i is the function of only one individual parameter X_i. Note that the definition above suggests an additive model is free from parameter interactions.

5.1.2 Connection with Standardized Regression Coefficients

It turns out that under certain circumstances, one will find a one-to-one correspondence between S_{X_i} and the squared standardized regression coefficients, as also mentioned by Mara and Tarantola (2008). To be more concise, for linear models with independent inputs, $S_{X_i} = \beta_{X_i}^2$. For instance, assume a linear model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$, where input parameters are normalized to mean 0 and standard deviation 1. It is easy to show that $S_{X_1} = \frac{V_{X_1}(E_{X_{-1}}(Y|X_1))}{V(Y)} = \frac{V(\beta_0+\beta_1 x_1+E(\beta_2 x_2+\epsilon))}{V(Y)} = \frac{\beta_1^2}{\beta_1^2+\beta_2^2} = \beta_1^2$. This result indicates that if our energy model is linear with regard to input parameters, we can compute first-order effect sensitivity index in variance-based methods with linear regressions.

5.2 Group Lasso with Discrete Categorical Variables and Sliced Latin Hypercube Sampling

Sensitivity analysis in our specific context desires the following features: it should enable a fair comparison between continuous variables (input parameters) and discrete categorical variables (indicators of different model forms), and be able to consider the combined sensitivity of several uncertainty parameters as a whole. We explain as follows a methodology that meets such requirements.

5.2.1 Consideration of Categorical Variables

We need to introduce categorical variables to include model forms in our SA. For instance, in the case HVAC uncertainty, the value of the respective categorical variable 0 denotes the model form that ignores HVAC uncertainty, while 1 indicates otherwise. Another use of categorical variables is for handling weather: if we have 41 years of historical weather data for Atlanta, then the categorical variable for weather will have 41 distinct levels, which can be translated to 40 dummy variables for regression purposes. The translation of multi-level categorical variables to dummy variables is explained in any statistical texts about regression.

The sensitivity index of the categorical variable indicates the importance of considering one type of model form uncertainty relative to other input parameters. However, the interpretation of the standardized regression coefficients should be altered to accommodate categorical variables, as the notation of mean or variance makes little sense for them. Hereby we choose to use $\frac{SSR-SSR_{-X_i}}{SSR}$ to measure the sensitivity of all parameters, discrete or continuous. In the equation above, SSR is the regression sum of squares for including all significant parameters from the group lasso, while SSR_{-X_i} denotes the regression sum of squares for leaving out parameter X_i from the pool. Therefore, the metric intuitively measures the reduced amount of variation explained by the regression model for leaving X_i out. For continuous independent parameters, $SSR = SSR_{-X_i} + \beta_i^2 V(X_i)$, so the measure is equivalent to previous notations in 5.1.2, but is now able to accommodate categorical variables.

5.2.2 Grouping of Variables with Group Lasso

Lasso (least absolute shrinkage and selection operator) is a regression method that conducts both variable selection and regularization, introduced by Tibshirani (1996). In its basic form, lasso's objective is to minimize (with regard to β)

$$\frac{1}{N}\|y - \boldsymbol{X}\,\beta\|_2^2 \text{ subject to } \|\beta\|_1 \leq t,$$

where $\|Z\|_p = (\sum_{i=1}^{N} |Z_i^p|)^{1/p}$ is the standard l^p norm. In addition, t is a tuning parameter, which determines the level of shrinkage. Let $\widehat{\beta^o}$ be the full least squares estimates and let $t_0 = \sum |\widehat{\beta^o}|$. If $t < t_0$, lasso will shrink all coefficients towards 0 by a similar amount, and sufficiently small coefficients are shrunken to exactly 0 (James, Witten, Hastie, & Tibshirani, 2013). Therefore, insignificant parameters are eliminated in the screening step.

Yuan and Lin (2006) propose group lasso as an augmentation to the basic version. Group lasso allows predefined groups of parameters to be selected in or out of the model

together. The most natural use of group lasso is that it can either include or exclude all levels of a categorical variable altogether, as it makes no sense to select only a few levels of a categorical variable. Suppose there are J groups of parameters, the group lasso estimation minimizes

$$\left\|y - \sum_{j=1}^{J} X_j \beta_j\right\|_2^2 + \lambda \sum_{j=1}^{J} \left\|\beta_j\right\|_{K_j},$$

where $\|z\|_{K_j} = (z^t K_j z)^{1/2}$, K_j is a positive definite matrix, and λ is a tuning parameter. In the formulation of group lasso, the design matrix X_j and coefficients β_j for each group of parameters replace the design matrix X and coefficients β in the old notation.

5.2.3 Sliced Latin Hypercube Sampling

Recall from Chapter 2 the basic Latin Hypercube sampling, which is a more efficient design than the traditional brute force Monte Carlo, for exploring the design space uniformly. However, now we are faced with the situation of sampling with the presence of categorical variables. If we use the traditional Latin Hypercube sampling, the design will not guarantee that at each level of the categorical variable, the remaining parameters fill the design space efficiently, which may give certain level of the categorical variable underserved advantage in the sensitivity analysis to come. Therefore, we advocate the application of the optimal sliced Latin hypercube sampling (SLHS) (Ba, Myers, & Brenneman, 2015) as an extension to the traditional LHS. An SLHS has the following features: (1) for the m samples under each level of the categorical variable, all marginal distributions on continuous variables are stratified into m strata with equal probability; (2) for all the $N = m \times t$ samples combined (t is the number of levels for the categorical variable), the marginal distributions on continuous variables are stratified into N strata with equal probability; (3) the samples achieve the maximum uniformity and space-filling property.

5.3 Case Study

5.3.1 Sensitivity Analysis Description

In this section, we conduct a sensitivity analysis with the proposed methodology above on the medium office building from the DOE reference building pool. The building has three floors, and each consists of a core zone and four perimeter zones. The system in the building includes a central packaged air conditioning unit with a gas furnace for heating and cooling, and variable air volume (VAV) terminal boxes with reheat for air distribution. In the sensitivity analysis, we study quantities of interest such as whole-building energy consumption in Atlanta. Input parameter uncertainties are drawn from the UQ repository described in Chapter 3, and we consider the following categorical variables summarized in Table 5.1.

Table 5.1 Categorical variables considered in the case study

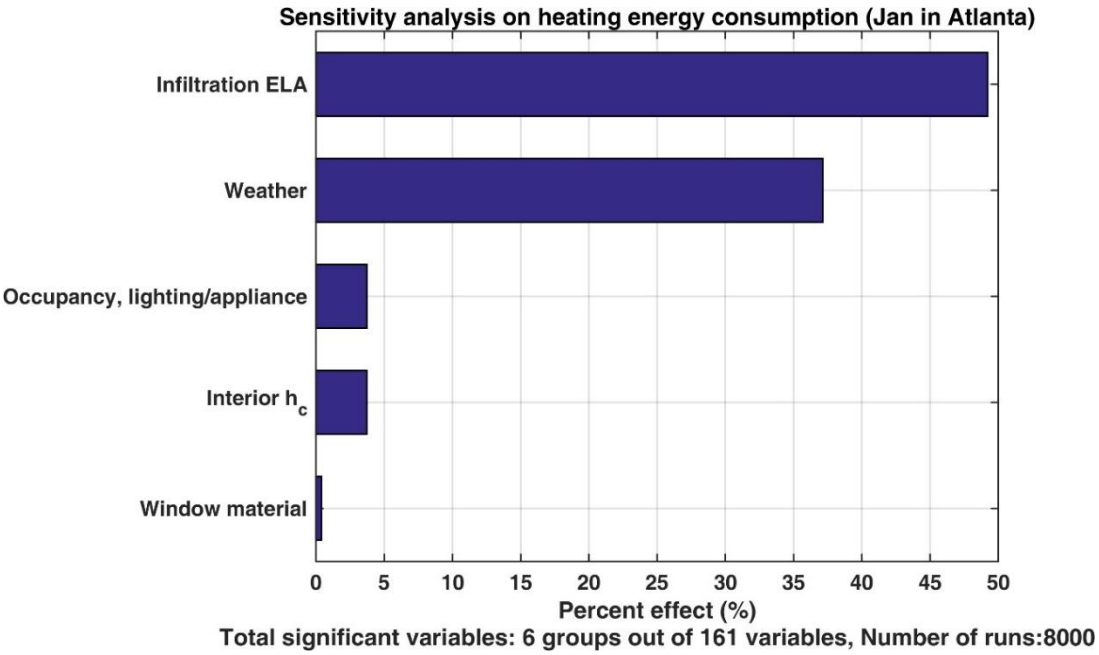
Historical Weather
Wind speed, wind pressure, infiltration
VAV systems
Occupancy, lighting/appliance usage

The first categorical variable deals with uncertainty associated with weather, as future weather is unknown at the design stage. We take 40 years of historical weather data in Atlanta to represent the variability of future weather conditions. For the second, we group wind speed, wind pressure and infiltration model form together, as they represent microclimate conditions around the building and they combine to affect the calculation of the amount of infiltration into the building. The third one is rated to HVAC systems. We implement quantified uncertainty with the EnergyPlus representation of VAV systems from Section 4.6. The last one is also from Chapter 4, with improved prediction fidelity of occupancy and lighting/appliance usage by a stochastic model. By turning on or off each of these model form uncertainties, we essentially derive $2^4 = 16$ distinct combinations of model forms.

In the sensitivity analysis, we first generate 16×500 samples with SLHS. Then, we combine the uncertainty analysis results at each level for the overall uncertainty distribution of the outcome of interest. In our implementation of group lasso, we use the well-established R package grplasso (Meier, 2015), with prescribed grouping indexes from our physical knowledge. For instance, all parameters on material properties are grouped as one, while those on the interior convective heat transfer coefficient are grouped as one. In total, we divide 161 parameters to 12 groups, i.e., ground reflectance, internal heat transfer coefficient, exterior heat transfer coefficient, ground temperature, wall materials, window materials, duct pressure loss, weather, wind pressure/speed, VAV model form, occupancy and lighting/appliance. These groups will be included or excluded altogether by lasso selection and regularization. Finally, sensitivity ranking of remaining significant parameters are ranked with the measure $\frac{SSR-SSR_{-X_i}}{SSR}$.

5.3.2 Results and Discussions

We show the results of sensitivity analysis on both heating and cooling energy for the case study building in Atlanta, in Figure 5.1.



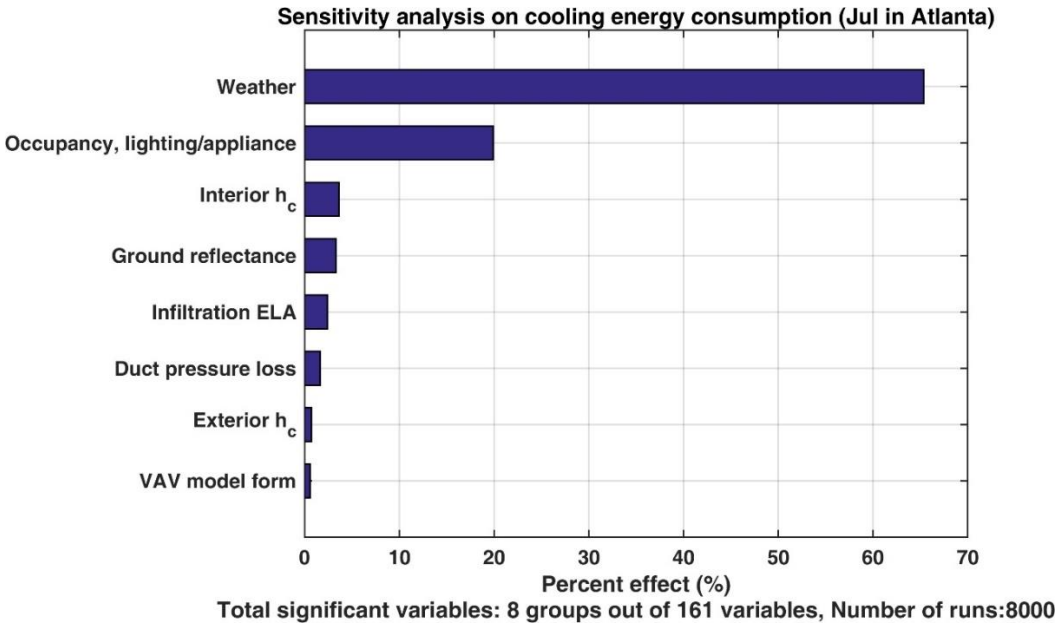


Figure 5.1 Sensitivity index for heating (upper) and cooling (lower) consumption in Atlanta

In Figure 5.1, both parameter and model form uncertainties are ranked together. For instance, the sensitivity of infiltration ELA on heating energy consumption is high, followed by the effect of considering historical weather as a "scenario uncertainty". In terms of cooling energy consumption, weather and occupancy variables, as scenario uncertainties, play a dominant role. A deeper examination of the impact of occupancy, lighting/appliance shows that for aggregated energy consumption prediction in an office building, more efforts should be focused on the estimation of the mean profiles of lighting/appliance usage, as knowledge of occupancy presence is of much less importance for energy consumption outcomes (Wang et al., 2016). Another interesting finding is that the uncertainty quantification of local wind speed, pressure and the infiltration model show negligible sensitivity on both heating and cooling energy consumption.

5.3.3 Iterative Uncertainty Analysis

As an illustration of the iterative uncertainty analysis methodology in Chapter 2, we re-examine the uncertainty of infiltration ELA, as sensitivity results that we just show rank ELA highest on the influential factors for heating energy. Currently, from the generic uncertainty quantification in Chapter 4, we characterize the distribution of ELA as $lognormal(1.282, 0.879^2)$, when no specialized information on the building of interest is available. However, to manage uncertainty and risk for decision-making, we assume a blower door test is performed shortly after the building is built, which greatly reduces the uncertainty of ELA. From the measurements, the ELA is estimated to be following the distribution of $normal(2, 0.1^2)$. Therefore, we are able to conduct a second-iteration uncertainty analysis (maybe also with sensitivity analysis) on energy outcomes, which enhances our confidence in the underlying prediction model for decision-making scenarios.

Figure 5.2 shows the comparison between predicted heating energy consumption with or without specialized information, considering all uncertainties besides weather. Note that the uncertainty range has been reduced significantly. The resulting model may infuse more confidence into certain decisions such as related to the need for operational intervention or confidence in fault detection.

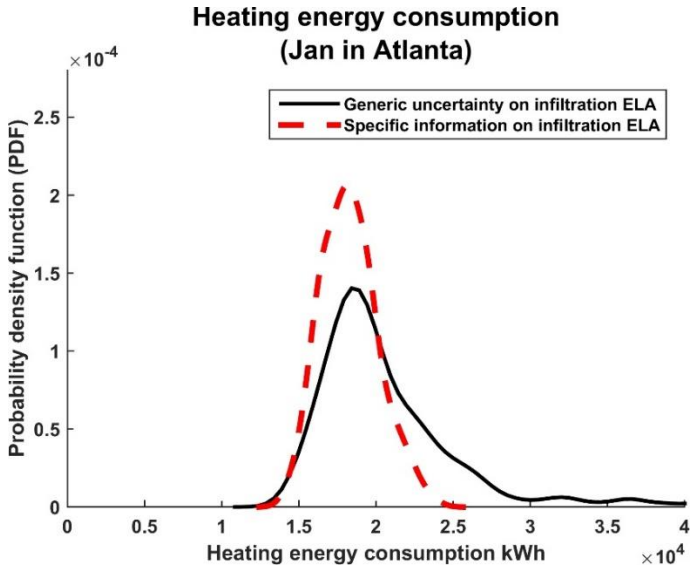


Figure 5.2 Result of original and refined uncertainty analysis

CHAPTER 6

VERIFICATION OF PROBABILISTIC ENERGY PREDICTIONS

The result of an uncertainty analysis is an expression of how accurate our model is, which itself requires verification or validation. Schlesinger introduces formal definitions of verification versus validation as the following: verification is the process of comparing computerized model with a conceptual model, while validation compares a computerized model with reality. Practically, both terms have been used interchangeably in different engineering domains (Roache, 2009). We choose the term "verification" in this research, following the norm in the climatology literature (Toth, Talagrand, Candille, & Zhu, 2003). In the building simulation domain, verification of probabilistic predictions has become an essential component of improving model accuracy.

6.1 Conventional Approach

The broader context of verification and validation of computer models can be found in Oberkampf and Roy (2010). Figure 12.1 of that work presents a variety of methods for comparing simulations and experimental measurements. Our application is most similar to Case e, except that our input is multi-dimensional (about hundreds) and systematic errors in the measurements are considered small compared to uncertainty associated with input parameters. In other words, we implicitly assume that the data of monthly energy use are of high accuracy. This assumption may hold because of the reputation of the installation company, the maturity of the measurement technology, and the aggregation use of the data.

Specifically in the building energy domain, work by Judkoff, Wortman, O'Doherty, and Burch (2008) summarizes the methodology used in the Building Energy Simulation Test (BESTEST) project by the National Renewable Energy Laboratory and

ANSI/ASHRAE Standard 140. The framework comprises of three parts: comparative studies, analytical verification and empirical validation. In a comparative study, results from two or more simulation tools are compared using equivalent input. If tools with different modeling approaches reach agreement, it is a strong indicator of the adequacy of these tools. Discrepancy, however, indicates cases where an empirical validation might be warranted. The analytical verification is straightforward, where simulation results are compared with simple test cases that can be solved analytically. This step eliminates the possibility of a major flaw in the heat transfer mechanism of tools, but does apply to a reasonably complicated case in reality. Empirical validation subjects outputs of a tool to the test by taking measurements in a real building or test cell. More often than not, the experiments are highly controlled, in the sense that input data acquisition is taken with great care to reduce uncertainty. Previous work defines three classes of control over error (uncertainty in our terminology) sources. Class A experiments isolate all sources of errors, but they typically only involve a few physical components or phenomena. Class B experiments control most sources of errors, and therefore they are typically only confined to simple test cell setups. Class C experiments do not control any error sources such that the experiment captures building operations in reality. The higher the class of control, it is easier to allocate remaining uncertainty and find out about unknown unknowns. In this dissertation, since we particularly deal with one tool and we are more interested in realistic rather than simple theoretical cases, we focus on the empirical validation approach.

Since the results of traditional computerized experiments are deterministic in nature, in a traditional validation context metrics such as the squares error and bias error are typically used. Suppose x_i is a simulation data point and y_i is a measurement, the squared error and bias error are calculated with $(x_i - y_i)^2$ and $(x_i - y_i)$. Typically arithmetic average is taken over hours or months, to compute the mean squared error (MSE) and mean bias error (MBE). For instance, ASHRAE 14 (ASHRAE, 2002)

stipulates that monthly simulation results have to be with 5% in terms of NMBE with regard to measurements, as an indication that the computer model has been validated.

Now we are ready to emphasize two distinct prerequisites for the verification of probabilistic building energy models to be introduced in this chapter. The first is the interpretation of measurement data: if we take as an example x_i , total energy consumption for a building in the month of July, it is important to realize that x_i is a random realization of the underlying random variable X. We issue probabilistic forecast F with cumulative distribution function (CDF) for X to describe the probability that X will take a value less than or equal to a given value. Since we cannot conduct a controlled experiment with a reasonably complex building, we are short of repetitive observations of the underlying X, but we realize that X is very likely to take on another value x_j governed by F, if there were an identical building built at the same location as the one of interest to us. The ramification of such a realization is that one will not immediately prefer the probabilistic prediction A on the left of Figure 6.1 to prediction B, without consideration of the likelihood of a true underlying distribution as also shown in the figure. If one considers this matter even further, he will also realize that the famous NBI data (Turner & Frankel, 2008) shown on the right of Figure 6.1 are not appropriate indicators of the notorious "performance gap", as the design EUIs should be associated with uncertainty while the diagonal line is only a comparison of the mean predicted EUIs against measured realizations.

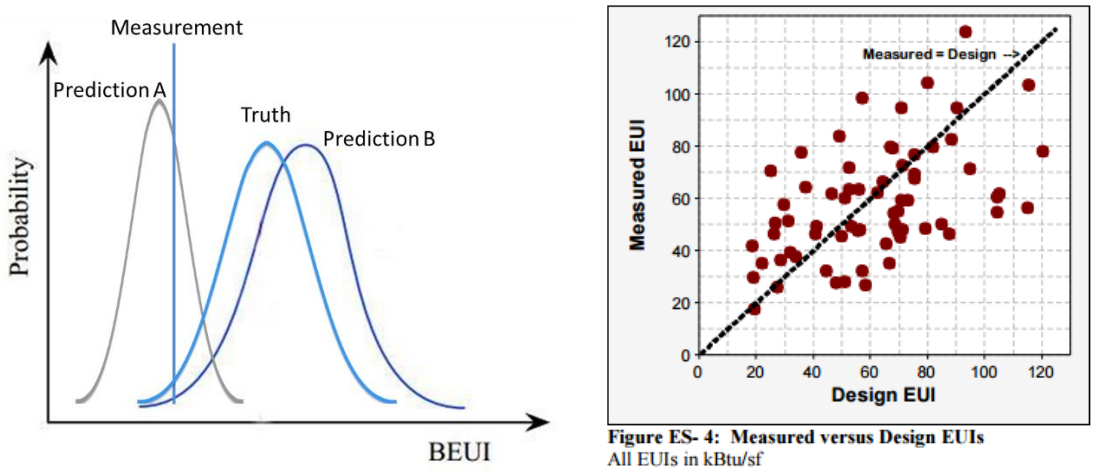


Figure ES- 4: Measured versus Design EUIs
All EUIs in kBtu/sf

Figure 6.1 True distribution versus measurements (left) and NBI data (right, adapted from Turner and Frankel (2008))

The second prerequisite is that a probabilistic forecast F evolves with the amount of information we possess when issuing it. Consequently, verification of a specific F cannot go without an explicit recognition of the uncertainty information that backs it up, for instance, whether the associated uncertainty quantification is rather generic or specific to the building of interest. The differentiation between specific and generic uncertainty quantification stems from the following philosophical conundrum: uncertainty quantification is by definition case specific, but at the design stage where this dissertation focuses on, modelers have no access to specific knowledge than generic.

In the remaining of this chapter, we detail proposed framework for verification of probabilistic building energy models, interpretation and simulation examples, and finally case studies applying this methodology to the estimation of an aggregated HVAC uncertainty factor for a pool of buildings under investigation.

6.2 Methodology

We propose the following methodology in the verification of probabilistic predictions for empirical validations. Traditionally, probability integral transform (PIT) (Rosenblatt, 1952) is used for testing whether a set of observations $x_i s$ can reasonably be

modeled as arising from a specified distribution Fx, which require multiple observations of the random variable. The complication in our application is that we measure only one realization from the underlying distribution that governs all possibilities for, say, the monthly total energy consumption of a building. Gneiting and Katzfuss (2014) propose the definition of a prediction space and generalize the traditional probabilistic integral transform.

6.2.1 Formal Definitions

Consider a real-valued observation Y, for which a probabilistic forecast F can be identified with associated cumulative distribution function (CDF) defined on the real line, $\mathbb{R}$. The prediction space setting considers the joint distribution of the probabilistic forecasts and the observations. The elements of the prediction space can be identified with tuples of the form (F, Y), where the probabilistic forecast F is a CDF-valued random quantity that utilizes a certain information basis or information set $\mathcal{A}$, which comprises of the training data, expertise, theories and assumptions. To be more concise, in measure theory, $\mathcal{A}$ is a sigma field, and F is a CDF-valued random quantity that is measurable with respect to $\mathcal{A}$. If we use $\mathcal{L}$ to denote an unconditional or conditional distribution, by definition, the CDF-value random quantity F is "ideal" relative to the information set $\mathcal{A}$ if $F = \mathcal{L}(Y|\mathcal{A})$. Essentially, an ideal probabilistic forecast makes the best possible use of information at hand. For example, suppose that $Y|\mu \sim N(\mu, 1)$ and $\mu \sim N(0,1)$. The probabilistic forecast $F = N(\mu, 1) = \mathcal{L}(Y|\mu)$ is ideal relative to the information set generated by the random variable μ. The forecast $F = N(0,2)$ is ideal relative to the trivial information set $\sigma(\emptyset) = \{\emptyset, \Omega)$, where Ω denotes the entire sample space.

Theorem 2.8 in Gneiting and Ranjan (2013) proves the following statement:

- A forecast that is ideal relative to a σ-algebra is both marginally calibrated and probabilistically calibrated.

The definitions of marginal and probabilistic calibration (calibration here means the statistical compatibility of probabilistic forecasts and observations) are as follows. In the prediction space setting, let F be CDF-valued random quantities with probability integral transform Z_F.

- The forecast F is marginally calibrated if $\mathbb{E}_{\mathbb{Q}}[F(y)] = \mathbb{Q}(Y \leq y)$ for all $y \in \mathbb{R}$.
- The forecast F is probabilistically calibrated if its probability integral transform Z_F is uniformly distributed on the unit interval.

Another technique, namely the reliability diagram (Wilks, 2011) is shown to be equivalent to checking for probabilistic calibration for a binary outcome (Gneiting & Ranjan, 2013). However, we need to realize marginal and probabilistic calibrations are necessary but not sufficient conditions for a forecast to be ideal, as Gneiting comments: "An interesting open question is whether there are any forecasts that are both probabilistically calibrated and marginally calibrated, but are not ideal. While we conjecture that the answer is in the positive, we do not know of any such examples". Therefore, we argue that probabilistic and marginal calibration are strong requirements for a forecast to be ideal. Checking for the both calibrations should form a cornerstone of density forecast evaluation.

6.2.2 Interpretations

A prediction space is used to model the joint distribution $\mathbb{Q}$ of the forecast F and observation Y. In other words, F is not a CDF as we traditionally have, instead it is a random variable taking the value of CDFs, and Y is a random variable taking the value of real numbers, and they have a joint distribution $\mathbb{Q}$. It might be difficult to understand exactly how F can be a random variable taking the value of CDFs. A naïve interpretation without seeking to measure theory is the following: suppose $F = N(\mu, 1)$ and $\mu \sim N(0,1)$,

then we realize F is random in a hierarchical sense: the CDF of F will be $\frac{1}{2}[1 + \text{erf}\left(\frac{x-\mu}{\sqrt{2}}\right)]$, but it is still random since μ is unknown.

As an illustration of the prediction space, we take the example of the verification of probabilistic predictions of hourly wind speed: since at each hour, the forecast issues a different prediction in terms of a CDF, and there is only one realization of wind speed each hour, the traditional PIT technique does not apply. However, the extended prediction space is able to accommodate this situation: at each hour, we have a realization from the joint distribution $\mathbb{Q}$ as a forecast-observation pair, where the forecast is a specific CDF and the observation is a real number. In practice, if we have multiple points in time (say 168) available for evaluation, that gives us vectors $(F_1, F_2, \dots, F_{168})$ and $(y_1, y_2, \dots, y_{168})$. If we use a time series model to predict hourly wind speed, for instance, we predict wind speed at hour t with wind speeds at previous two hours $Y_t = f(Y_{t-1}, Y_{t-2}) + \varepsilon$, then the smallest information set in this case will be the σ-algebra $\mathcal{A}_t = \sigma(Y_{a-1}, Y_{a-2})$ generated by the two necessary past observations.

An ideal probabilistic forecast relative to a specific information set makes the most efficient use of the information set. To check probabilistic calibration of forecast, we plot a histogram of the vector $F_1(y_1), F_2(y_2), \dots, F_{168}(y_{168})$ and see if they are uniformly distributed on the unit interval. For marginal calibration, we check whether the empirical CDF of Y corresponds to $\frac{1}{168}\sum_{i=1}^{168} F_i\,(y)$ for all $y \in \mathbb{R}$.

As such, we are ready to apply the above theory to the verification of probabilistic building energy models. The analogy requires that each forecast-observation pair we get from a building should be a random draw from the joint distribution $\mathbb{Q}$ of (F, Y). This stipulates that Y has to originate from the underlying data generating process for all similar buildings under investigation. In our application, the information set is then generated by (shape, materials, usage scenarios, physical principles, uncertainty propagation assumptions, etc.), or $\mathcal{A} = \sigma(X_1, X_2, X_3)$. An implicit assumption is that

similar combinations of values for shape, materials, etc. will lead to similar forecasts and similar realizations. Although we assume forecast-observations pairs from similar building under investigation are random draws from the same joint distribution $\mathbb{Q}(F, Y)$, the conditional joint distributions $F_i = \mathcal{L}(Y|(x_1, x_2, x_3))$ for specific buildings may of course differ depending on the value of shape, materials, etc. We will present a simulation example later this chapter to make this idea more concrete.

6.2.3 Diagnosis for Probabilistic Calibration

If we plot the empirical CDF of PIT values against theoretical CDF of a uniform distribution, the discrepancy between them will indicate reasons for forecast deficiency. For instance, Sun (2014) displays with simulation examples, forecasts that are under-dispersed, over-dispersed, biased due to underestimate, and biased due to overestimate. For an over-dispersed forecast, a sensitivity analysis may be conducted, and the uncertainties contributing most to the spread of outcome may need a re-examination. If the predictive distributions are under-dispersed, we should be looking for potential significant uncertainties neglected in our quantification, realizing there are unknown unknowns in our knowledge.

6.2.4 Ranking of Competing Forecasts

When we know that two competing forecasts are both ideal, and their information sets are nested, the forecast with the larger information set should be ranked higher. Otherwise, a proper scoring rule, such as the continuous rank probability score (CRPS) (Matheson & Winkler, 1976) should be used. CRPS is given by the following equation:

$$CRPS = \frac{1}{ncases} \sum_{i=1}^{ncases} \int_{-\infty}^{\infty} (F_i^f(x) - F_i^0(x))^2 \, dx,$$

where

$ncases$ is the number of forecast cases,

$F_i^f(x)$ is the forecast CDF for the ith forecast case,

$F_i^0(x)$ is the observation, expressed as a CDF, for instance, if the observation is a single value, then the corresponding CDF is a single step-function with the step from 0 to 1 at the observed value of the variable.

In Figure 6.2, we show the CRPS the shaded area, for a case where the observation is a single value.

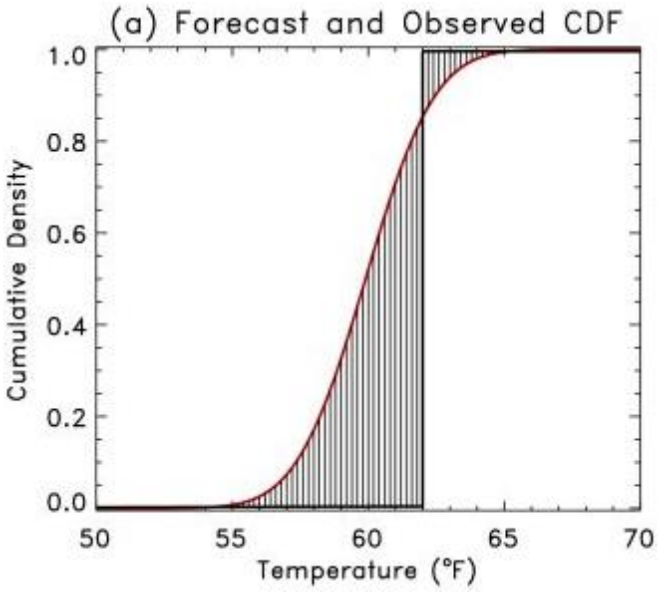


Figure 6.2 CRPS for a single observation

6.2.5 Simulation Examples

In this section, we provide simulation examples for clarifying our methodology further. Suppose the true data generating process for each building in our dataset is assumed to be $Y = X_0 + a_1X_1 + a_2X_2 + \varepsilon$, where $X_0, X_1, X_2 \sim N(2,1)$, $a_1 = a_2 = 2$ and $\varepsilon \sim N(0,1)$ for simplicity. These random variables could be interpreted as shape, materials, usage scenario, etc. and the linear formulation and ε represent assumptions on the physical principle and uncertainty quantification. It is easy to see that forecast $F_1 = N(X_0 + a_1X_1 + a_2X_2, 1)$ is not only ideal with respect to the information set $\sigma(X_0, X_1, X_2)$, but F_1 is the perfect forecast since we have all the information needed by the data generating process. In this situation, a particular building j is a forecast-

observation pair drawn from (F, Y), and $F_{1,j} = \mathcal{L}(Y|(x_{0,j}, x_{1,j}, x_{2,j}))$ can be interpreted as the conditional distribution of Y given a certain realization of random variables (X_0, X_1, X_2). A less informed forecast can be $F_2 = N(X_0 + a_1 X_1 + 2 \cdot a_2, 1 + a_2^2)$, which does not have access to private information about X_2. However, it is still ideal with respect to a smaller information set $\sigma(X_0, X_1)$. If we can show that both forecasts are ideal by checking probabilistic and marginal calibration, F_1 should be ranked higher since its information set is larger. In addition, it is expected that the variance of F_2 is larger for a lack of knowledge.

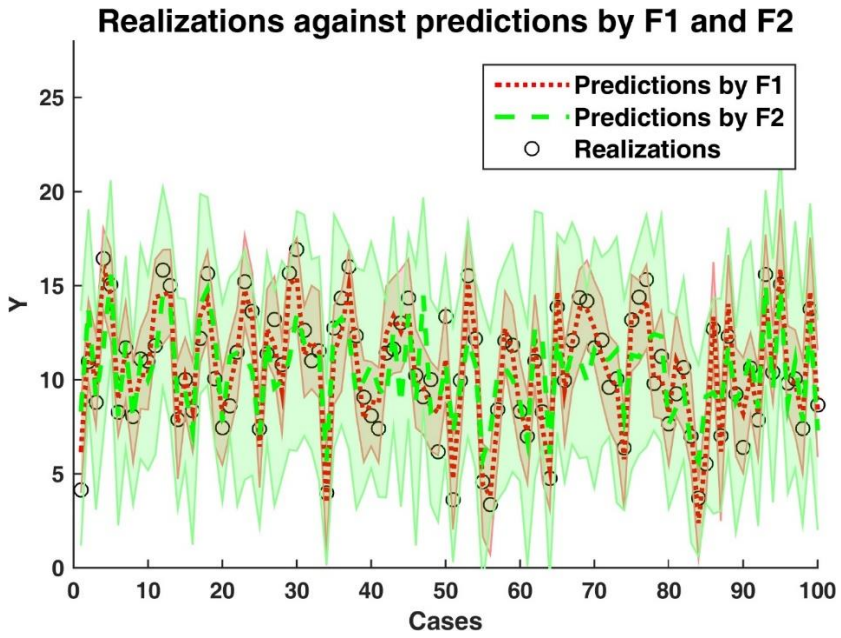


Figure 6.3 Realizations against prediction by both forecasts

In Figure 6.3, we show the realizations against predictions by both forecasts: dotted lines denote the mean predictions, and the shaded areas represent the uncertainty ranges. We can see that the predicted range by F_2 is much wider than that by F_1 for a lack of information. However, when we subject both forecasts to test by probabilistic calibration and marginal calibration, we find interesting results. In Figure 6.4 left, we compare empirical CDF of probability integral transforms with the theoretical one that is uniformly distributed on the unit interval, and we show on the right the expected CDF

from forecast predictions and the empirical CDF of realizations. The fact that these curves overlap suggests that both forecasts are ideal relative to its own information set. In this situation, we conclude that both forecasts are ideal, but F_1 is sharper (Gneiting, Balabdaoui, & Raftery, 2007) and should rank higher than F_2.

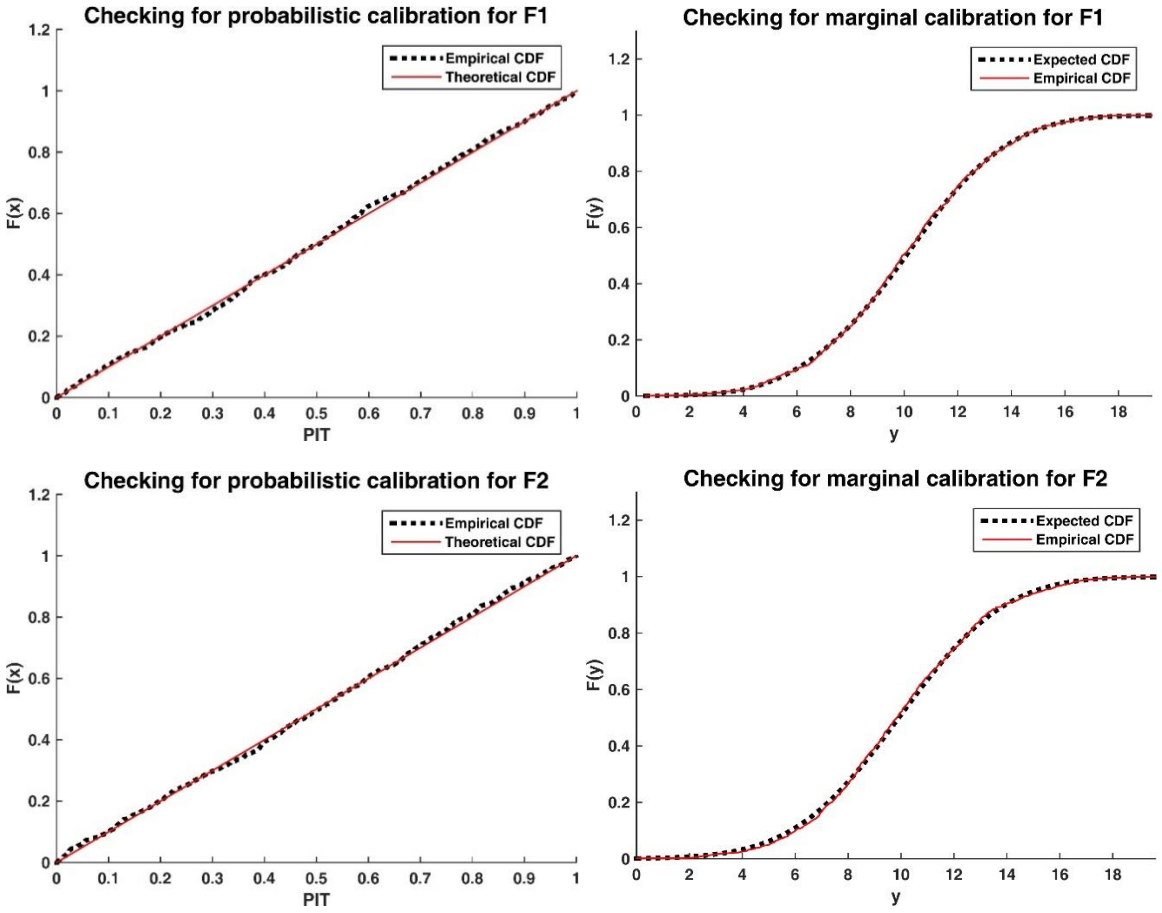


Figure 6.4 Probabilistic (upper) and marginal (lower) calibration results by forecasts

6.2.6 HVAC Uncertainty Factor

Realizing the existence of many unknowns about existing HVAC systems, for instance, maintenance practice, faults, etc., we argue that rudimentary calibration to our design stage model makes little sense. As such, we switch to a top-down approach and lump all HVAC model form uncertainties and the impact of faults into an overall HVAC uncertainty factor. To elaborate, the top-down approach in this chapter will first model

six selected campus buildings as truthfully as possible, and propagate all input parameters, model form and scenario uncertainties into the outcome. The resulting probabilistic predictions will be compared against measurement data. At this stage, we expect the probabilistic forecast will fail the probabilistic calibration and marginal calibration, resulting from a lack of consideration of unknown uncertainties. We then superimpose an HVAC uncertainty factor onto the previous probabilistic predictions and conduct a parametric analysis, until both tests pass. The resulting uncertainty factor will be interpreted as mainly caused by the ensemble of HVAC idealization and system faults that are likely to occur in the buildings we are investigating. The finding will provide guidance to the study of future similar buildings, as an estimation of the magnitude of overall HVAC uncertainty.

6.2.6.1 Case Study Buildings

We choose to study six buildings on campus at Georgia Institute of Technology for a demonstration of the methodology proposed above. Table 6.1 summarizes general information about these buildings. What in common between them is that they are all mixed-use buildings, mainly comprised of offices, classrooms, and labs. All buildings have HVAC systems for maintaining indoor thermal comfort, and the cooling and heating sources are from local systems, instead of district systems.

Table 6.1 Building general information summary

Name	Space types	Gross floor area (m^2)
Georgia Tech alumni association	Office, accounting	2,362
Office of human resources	Office	1,510
Ivan Allen college of liberal arts	Office, classroom	2,709
Georgia center for advanced telecommunications technology (GCATT)	Office, software lab, hardware lab, conference room, auditorium	14,628
Language institute O'Keefe building	Office, classroom, library	10,224
Institute of paper science and technology (IPST)	Laboratory, classroom, office, conference room, auditorium, library,	15,136

The Georgia Tech alumni association has three floors and its floor area is 2,362 m^2. Variable flow AHUs connected with terminal VAV boxes supply conditioned air to each space. Heating and cooling are supplied by an air-cooled chiller and a boiler. The office of human resources is a two-story building with a floor area of 1,510 m^2. A rooftop unit connected with fan powered terminal units (with electric reheat capability) provide heating and cooling to the building. The Ivan Allen college of liberal arts building has four floors with a floor area of 2,709. The building is conditioned by local split system heat pump, for both cooling and heating. The Georgia center for advanced telecommunications technology is a five-story building with a floor area of 14,628 m^2. The HVAC system mainly consists of variable flow AHUs connected with fan powered

terminal units (FPTU) or VAV boxes at each branch. There are also a few fan coil units and unit ventilators. A central cooling and heating plant with chillers and boilers supplies chilled and hot water to AHUs and reheat coils at FPTUs. Heat from the building is rejected to the environment via cooling towers. The Language institute O'Keefe building is a three-story building with a floor area of 10,224 m^2. The HVAC system type is the same as the Georgia center for advanced telecommunications technology, except that there are no fan foil units. The Institute of paper science and technology building is a five-story building with a floor area of 15,136 m^2. The HVAC system is similar to Georgia center for advanced telecommunications technology, but one specialty is that there are a lot more exhaust fans installed tailored to the needs of several labs.

We study detailed architectural and mechanical drawings for all six buildings, based on which we model the building geometry, construction types, thermal zoning, and HVAC systems as truthfully as possible. We propagate input parameter uncertainty, model form and scenario uncertainty through our GURA-W workbench, and produce probabilistic predictions. For the six buildings under investigation, we collect actual weather conditions and site energy consumption data for each month from 2010 to 2013. Site energy consumption data are total energy consumed each month in terms of kWh, including electricity and natural gas for most buildings, as a detailed separation between them is not available. Consumption data are then normalized by the building total floor area to derive energy use intensity, respectively.

6.2.6.2 Preliminary Forecast Predictions Versus Measurement Data

In the following figures, we present a visual comparison between monthly probabilistic energy forecasts and measurement data. For computational simplicity, we choose one typical year out of measured four years for each building in the comparison. In these figures, the green dots represent measurement data, while the boxplots represent predicted uncertainty range for monthly energy consumption.

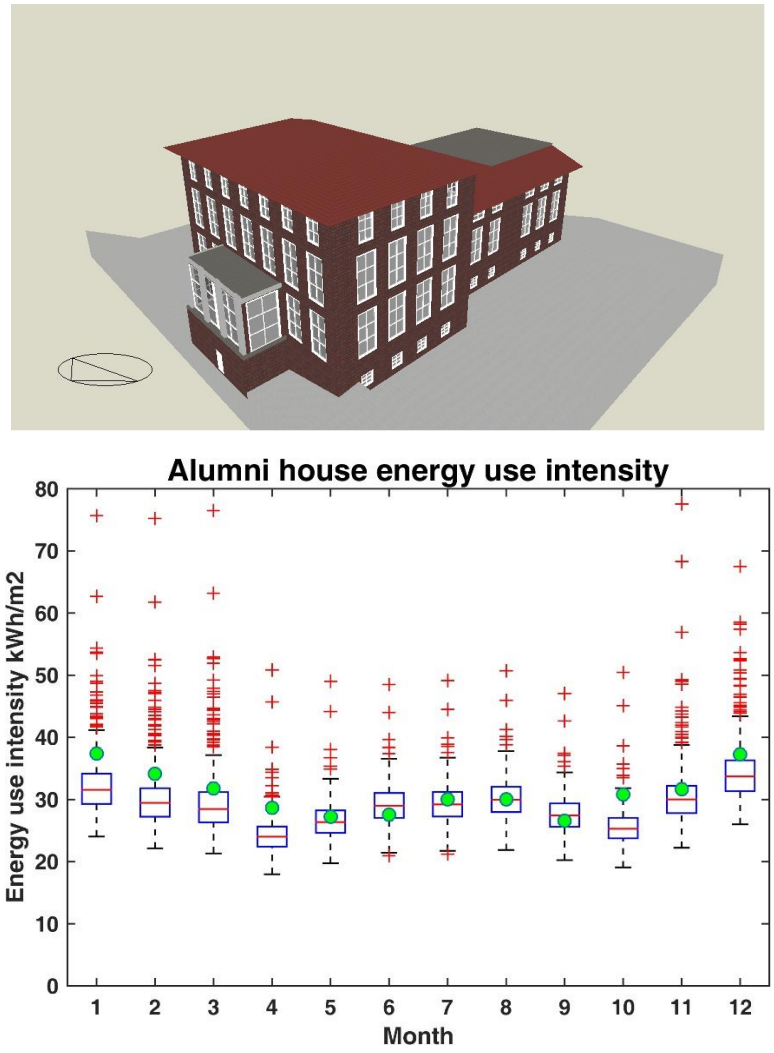
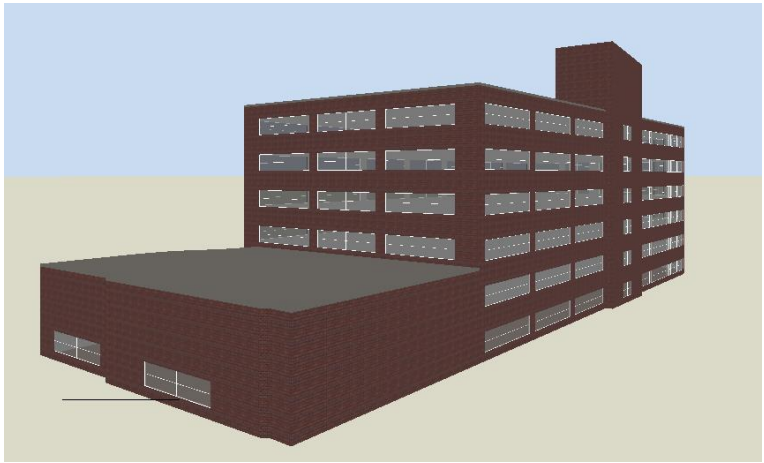


Figure 6.5 Georgia Tech alumni association



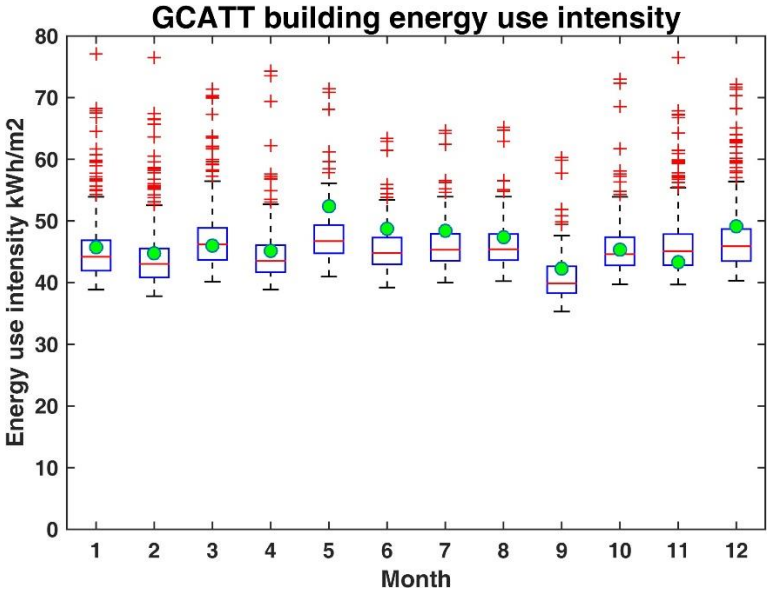


Figure 6.6 Georgia center for advanced telecommunications technology (GCATT)

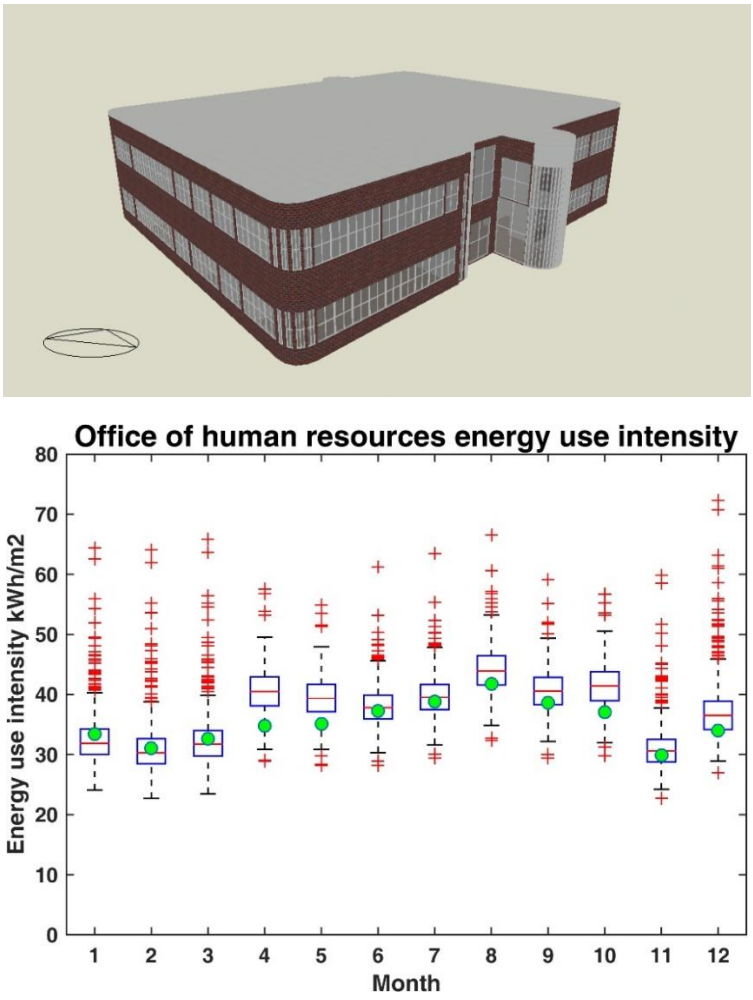


Figure 6.7 Office of human resources

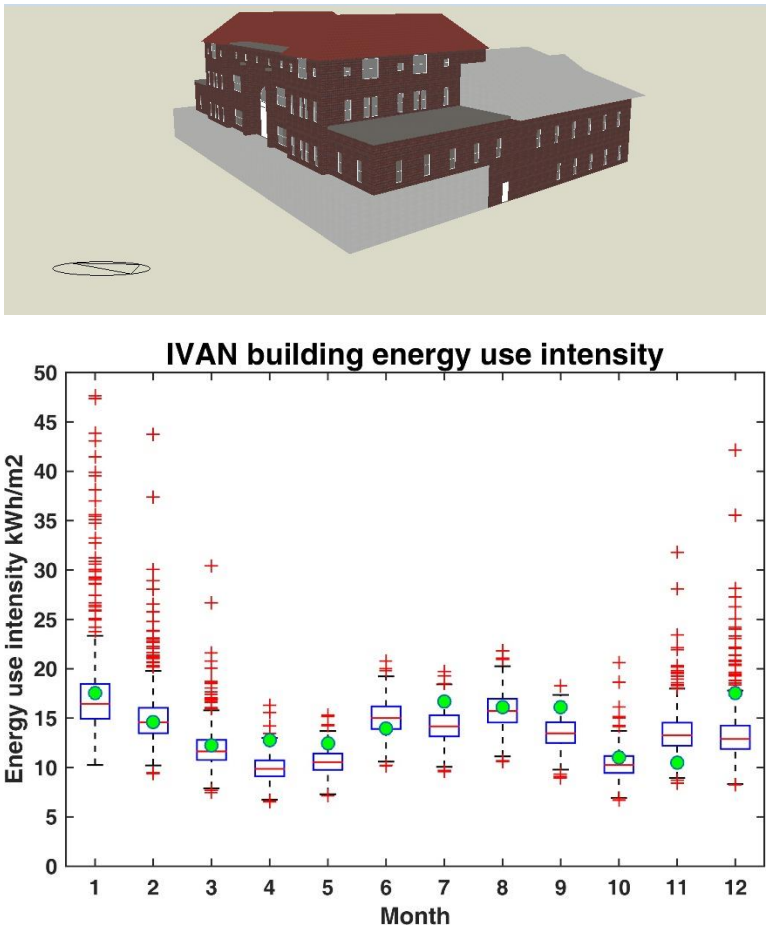
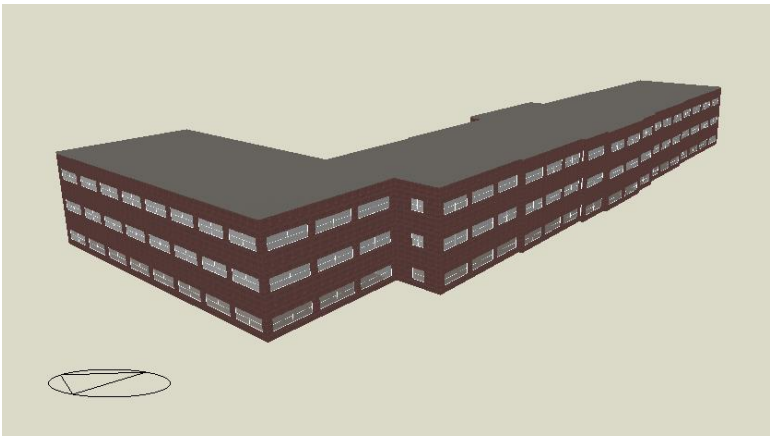


Figure 6.8 Ivan Allen college of liberal arts



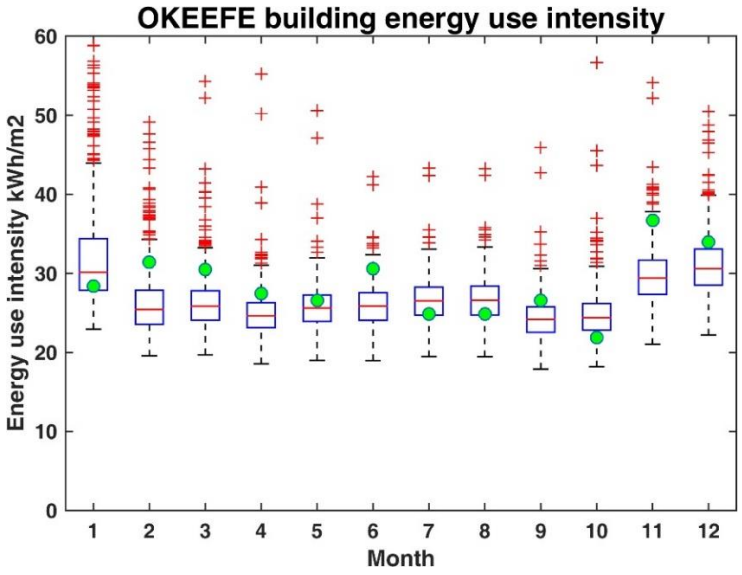
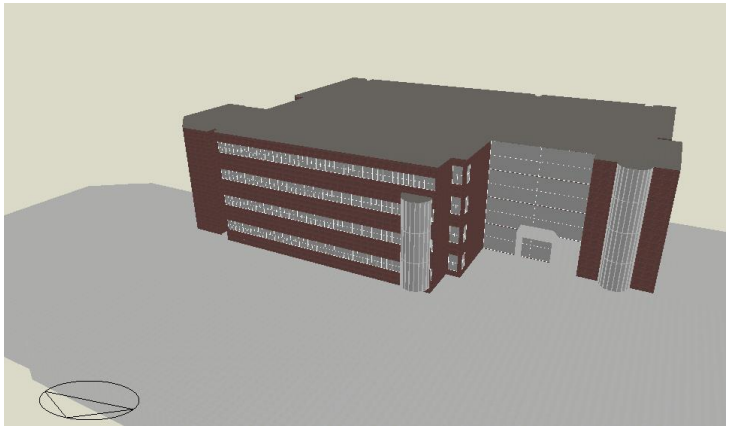


Figure 6.9 Language institute O'Keefe building



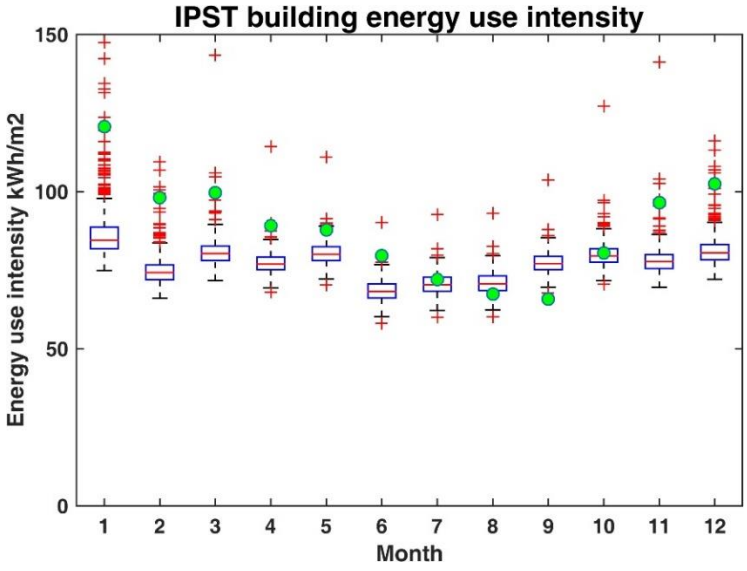


Figure 6.10 Institute of paper science and technology (IPST)

As we can see in the figures above, almost all measurement data points fall within the uncertainty range (represented by the box plots) predicted by the forecast. The red points on the figures denote outliers that we expect result from the parameter of ELA that is lognormally distributed. An explanatory analysis of the measured energy use intensity for all six buildings reveal that the consumption of the Institute of paper science and technology building is much higher than the rest of the buildings. Heavier electrical loads for laboratory use in IPST building may explain why the forecast with generic usage uncertainty under-estimates the building energy use intensity. Consequently, this may violate the underlying assumption of our whole approach that we can only pool buildings with similar combinations of values for materials, usage, etc. as they lead to similar forecasts and realizations. We will re-examine this issue in the next section of probabilistic and marginal calibration.

6.2.6.3 Preliminary Test of Probabilistic Calibration and Marginal Calibration

Pooling the monthly energy consumption data from all six buildings together, we have in total 72 data points (or cases). For each month, we have a forecast-realization pair that is drawn randomly from the underlying joint distribution $\mathbb{Q}(F, Y)$. To check probabilistic calibration of forecast, we plot a histogram of the vector $F_1(y_1), F_2(y_2), \ldots, F_{72}(y_{72})$ and see if they are uniformly distributed on the unit interval. For marginal calibration, we check whether the empirical CDF of all monthly energy use intensity corresponds to $\frac{1}{72}\sum_{i=1}^{72} F_i\,(y)$ for all $y \in \mathbb{R}$.

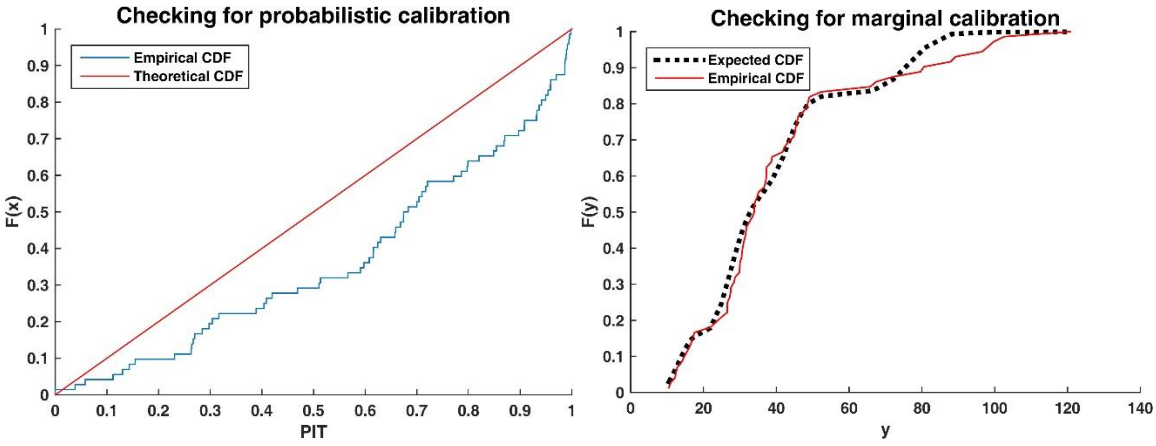


Figure 6.11 Preliminary calibration results: probabilistic (left) and marginal (right)

Figures 6.11 shows results for the probabilistic calibration and marginal calibration. Theoretical curves and empirical ones are not quite far apart, which indicates preliminary modeling and uncertainty analysis have captured the basic characteristics of the energy consumption in the six buildings. From the diagnosis technique mentioned above, we realize that the preliminary forecast is biased due to underestimate. Besides, the discrepancy between expected and empirical CDF for marginal calibration indicates that for energy use intensity higher than 85kWh/m^2, the expected CDF value is always 1, indicating that high EUI is not expected by the forecast in general. This finding supports our previous concern that the IPST building with exceptional high EUI relative to others will contaminate our study. Therefore, we decide to exclude the IPST building, ensuring the underlying assumption remains intact: we can only pool buildings with similar combinations of values for materials, usage, etc. as they lead to similar forecasts and realizations. Figures 6.12 shows calibration results for the rest of the five buildings.

Improvements over the previous calibration results by removing the IPST building from the dataset can be observed easily, comparing Figure 6.11 and 6.12. In order to quantify the discrepancy between empirical and theoretical CDFs, we use the one-sided K-S test (ref) statistic, which in a discrete form can be represented by $D_n = \sup |F_n(t) - F(t)|$, in which t denotes the point at which the CDF is evaluated. The

statistic essentially measures the largest vertical discrepancy (absolute) between the CDF being evaluated and the baseline one. With this measure, we quantify a reduction of K-S statistic from 0.2574 (PIT) and 0.1085 (marginal), to 0.2240 (PIT) and 0.0904 (marginal). However, the K-S test for PIT rejects the null hypothesis that the CDF being evaluated is equal to the baseline CDF at the 5% significance level. This result indicates room for improvement over our preliminary forecast.

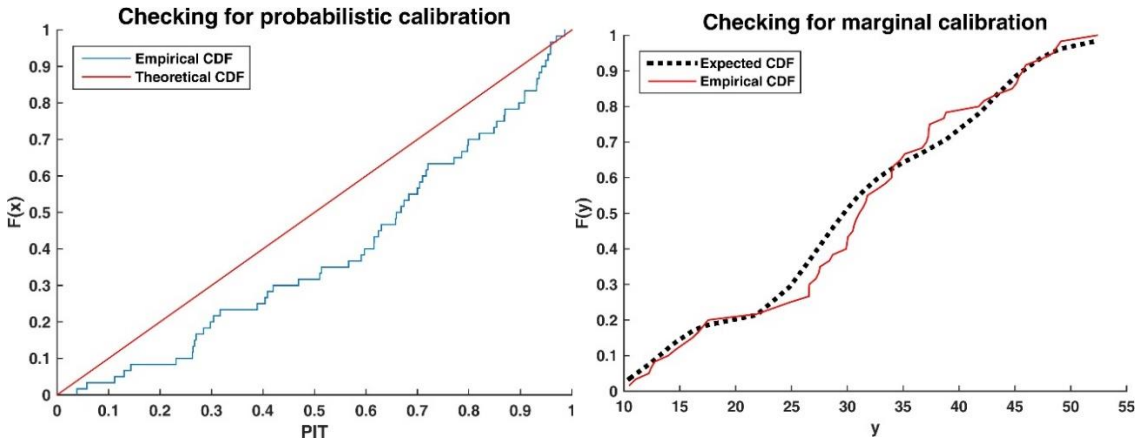


Figure 6.12 Calibration results without IPST: probabilistic (left) and marginal (right)

6.2.6.3 HVAC Uncertainty Factor

The reasoning behind an aggregated HVAC uncertainty factor is that we can attribute the deficiency of our preliminary forecast to a combined effect of modeling idealization and faults. This of course is premised on the loyalty of the simulation models to design drawings and a due diligence about the quantification of all other sources of uncertainty. We assume the HVAC factor for heating η_{heat} and η_{cool} each follows a normal distribution with mean μ and standard deviation σ. For a given building, we generate one realization of η_{heat} and η_{cool} respectively for each simulation run. Then we multiply the realization of η_{cool} with HVAC consumption in months between May and September and η_{heat} with HVAC consumption in rest of the months that may require heating: $HVAC_{heat,W} = \eta_{heat} \cdot HVAC_{heat,WO}$, and $HVAC_{cool,W} = \eta_{cool} \cdot HVAC_{cool,WO}$. The treated HVAC consumption, combined with original lighting/appliance consumption,

leads to updated predictions of whole-building energy consumption to be compared with measurement data: $EUI_W = Light/app + HVAC_{heat,W}$ $(or\ HVAC_{cool,W})$. We quantify the outcome of probabilistic and marginal calibration with the sum of K-S statistic from each calibration test: $ksstat_{overall} = ksstat_{PIT} + ksstat_{marginal}$. In summary, for each pair of $(\mu_{heat}, \sigma_{heat})$ and $(\mu_{cool}, \sigma_{cool})$ value, we apply the procedure described above and obtain the calibration metric $ksstat_{overall}$. This leads to a well-formulated optimization problem to find the best value combination of $(\mu_{heat}, \sigma_{heat})$ and $(\mu_{cool}, \sigma_{cool})$, which leads to the minimum discrepancy described by $ksstat_{overall}$ between forecast and desired CDF, with consideration of an aggregated HVAC uncertainty factor.

7000 optimization runs with the Darwin genetic algorithm leads to the following best design: $\eta_{heat} \sim N(0.996, 0.19^2)$ and $\eta_{cool} \sim N(1.157, 0.0754^2)$, with a minimum $ksstat_{overall}$ of 0.13, compared to 0.314 before optimization. The resulting probabilistic and marginal calibration is shown in Figure 6.13. The K-S test for PIT cannot reject the null hypothesis that the CDF being evaluated is equal to the baseline CDF at the 5% significance level.

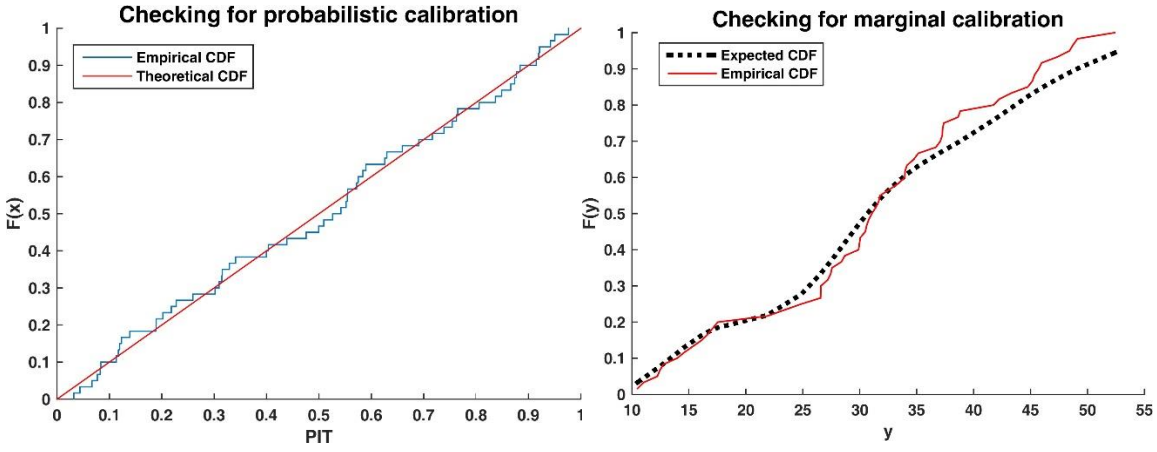


Figure 6.13 Calibration results with HVAC factor: probabilistic (left) and marginal (right)

Recall that the HVAC uncertainty factor subsumes the effect of HVAC modeling idealization and faults in real building operations. With the resultant η_{heat} and η_{cool}, we

estimate that our preliminary forecast is biased for under-estimation for cooling consumption by $1 - \frac{1}{\mu_{cool}} = 14\%$ with regard to measurement data. Furthermore, we find that although the forecast for heating consumption is not biased, enlarging the uncertainty range for HVAC consumption helps bridge the gap between predictions and measurement data. Another interesting finding is that the realization of η_{heat} or η_{cool} can be smaller than one, indicating the actual HVAC system consumes less energy than forecast predictions. This is no surprising if we realize that the HVAC system in reality may fail to maintain desired indoor temperature of air quality (e.g., insufficient outdoor air is supplied resulting from OA damper faults) whereas our prediction model in this case always assumes ideal fulfilment of temperature set points and design fresh air volume.

CHAPTER 7

VALIDITY OF PROBABILISTIC ENERGY PREDICTIONS

As we interpret validity of probabilistic predictions as leading to "better" decisions, a brief discussion on the definition of "better" is warranted. In this preliminary attempt, we define the best decision as the optimal decision computed by normative utility theory backed up by two key ingredients: carefully elicited risk preferences that reflect decision maker's value, and transparent uncertainty information for energy savings, from an ideal forecast verified by theory proposed in Chapter 6. For rigorous validity proof, we have to run experiments with decision makers that use uncertainty analysis and see how confident they are in making decisions. For time and resource constrain, we will just present one test case where we believe uncertainty and risk analysis add value to the decision making process.

7.1 Contractual Decisions in a Performance Contract

In order to demonstrate the impact of uncertainty quantification on decisions, we zoom into a specific context of energy efficiency investments, executed through performance contracting, where risks play a crucial rule. An energy savings performance contract allows building owners to upgrade their building with typically no up-front cost while repaying the project cost with the energy cost savings over the term of the contract. In a typical set up, an ESCO (Energy Service Company) either assists the client in arranging a loan from a third party investor, or finances the cost of developing, implementing and maintaining the energy conservation measures by itself.

The unique feature of a performance contract is that performance is "guaranteed"; the ESCO guarantees a specified level of cost savings for the client, which will be sufficient to finance the full cost of the project and will be verified by post-

implementation measurement and verification (M&V) activities. In a performance contract, there are typically four types of savings: (1) “projected saving” is an estimate of savings based on building energy auditing and a detailed energy study (probably conducted with building simulation software) by the ESCO; (2) “proposed saving” is the amount of savings that the ESCO is confident with and willing to claim as yearly “guaranteed saving”, i.e. the regular payment amount the client is expected to pay annually; (3) “verified saving” is the difference between the (adjusted) baseline utility cost and post-retrofit utility bill, calculated by the ESCO after annual M&V as agreed upon; (4) “actual saving” is the original amount of budget reserved for the cost of operation less the post-retrofit actual expenditure, which could be retained for other purposes. To elaborate, a baseline model is a statistical or physical model that describes the facility and its operations during the pre-retrofit phase. Note that this model can be adjusted for the post-retrofit actual weather condition and usage of the building to represent the future consumption of the building as if retrofits were not performed. Furthermore, a simple approach to distinguishing between verified and actual saving is that the former should only be used to determine whether the ESCO is subject to payout during M&V, while the latter should be used for calculating NPV or other financial measures for the building owner. In summary, performance contracting is meant to spread responsibility and risk over different entities. It is attractive to clients as they shift certain performance risk to the ESCO and they pay for nothing else (i.e., neither project implementation nor maintenance cost) but the future verified savings, which is based on the agreed upon M&V procedure. However, attention should be drawn to the point that in a performance contract, verified savings can substantially deviate from actual savings. This means that clients are not risk free but carry certain risks, such as the risk of extreme weather, energy price fluctuations, and the risk that certain contract stipulated clauses turn out to have negative consequences, whereas the ESCO carries the performance risk as well as the cost overrun risk.

As such, we realize both the client and ESCO can be subject to substantial risks in a performance contract. In the rest of this chapter, we first describe the normative utility theory as a tool for computing the optimal decisions that will rationally maximize the decision maker's value, and then we demonstrate how transparent risk information informed by uncertainty analysis can affect decisions made by the decision maker.

7.1.1 Axiomatic Utility Theory

According to normative utility theory, preferences over outcomes can be expressed in terms of utilities, the properties of which are outlined by the axioms of the theory. The key result of the theory is the expected utility theorem, which states that a decision maker should select the alternative with the greatest expected utility. The axiomatic foundation of utility theory was originally developed by von Neumann and Morgenstern (Neumann & Morgenstern, 1944). Furthermore, since one decision maker is more risk averse than the other, it is not surprising that two decision makers arrive at different choices in the same situation, which is solely the result of different risk preferences. Howard (1988) proposes a formulation of modeling risk preferences assuming constant risk tolerance. Under such an assumption, the approximating utility function is defined by $\mathrm{U}(\mathrm{wealth}) = \mathrm{R} \cdot \left(1 - \mathrm{e}^{-\mathrm{wealth}/\mathrm{R}}\right)$, where R is the constant risk tolerance and has the same unit as wealth.

7.1.2 Impact of Risk Information on Decisions

We illustrate the impact of risk information backed up by uncertainty quantification with the following real life project. The Anonymous middle school project is an energy retrofit project developed by ESCOMP[1] offering a comprehensive way of

[1] Due to a non-disclosure agreement, the name of the actual ESCO involved in the performance contract with the Anonymous school and the exact details of the proposed contract are not included. Instead, the ESCO is referred to as ESCOMP, and all data have been slightly altered to maintain the privacy of the ESCO.

updating a middle school building located in Pennsylvania. The building has a total floor area of 105,200 ft^2, and it serves approximately 500 students. Being in Pennsylvania, the school's HVAC operation is heating dominated. The previous yearly utility bill shows that the building consumes roughly 55,000 Therms/year of natural gas for heating and about 790,000 kWh/year of electricity for cooling and other uses. The current utility rate structure is electricity consumption at $0.082/kWh, electricity demand charge at $3.63/kW and natural gas consumption at $0.785/therm. After an investment grade auditing and a detailed energy analysis with simulation models, ESCOMP proposes the following ECMs:

- ECM1: Lighting retrofit,
    - Replace current T12 lamps with T8 lamps and electronic ballast,
- ECM2: Lighting control with occupancy sensor,
- ECM3: HVAC setback control,
    - Reduce heating setpoint during unoccupied hours at night,
- ECM4: Demand control ventilation,
    - Adjust fresh air supply in spaces like gymnasium and auditorium based on occupancy,
- ECM5: High efficiency boiler,
    - Replace the old boiler with a condensing boiler.

The following is an examination of the project with our proposed framework. First, to estimate the potential savings to be gained from the retrofit, we build an EnergyPlus model that matches the actual operational policy of the school observed and recorded on-site. However, uncertainties remain from our incomplete knowledge of the building, such as system parameters, construction detailing and workmanship issues, as well as weather. We combine uncertainties specific to the school building (detailed below) and generic model uncertainties specified in the UQ repository in the GURA-W.

- Condensing boiler efficiency: The efficiency of condensing hot water boilers depends on the return water temperature, so the constant efficiency coefficient typically used in simulations is uncertain. In addition, nameplate information with respect to efficiency will have manufacturing tolerance associated. In our study, we capture both effects as uncertainty in the efficiency coefficient (assuming a normal distribution with mean 0.904 and standard deviation 0.011) and a performance curve that correlates efficiency with the return water temperature (Lazzarin, 2012).
- Lighting consumption: National Lighting Product Information Program (NLPIP) (1999) identifies that the nameplate power of many lighting fixtures generally deviates from field tested values, so we introduce an uncertainty distribution of the percentage difference between monitored and nameplate power by fitting to multiple test samples reported by NLPIP to reflect the variability. We find that a reduction factor $L \sim N(0.95, 0.09)$ should be multiplied with the nameplate power to calculate aggregated consumption of T8 fixtures.
- Weather: To account for the variability in the weather, the approach developed by Lee (2014) has been used to generate stochastic meteorological year samples based on 33 years of historical meteorological data for Pittsburgh, Pennsylvania.
- Other uncertainties include those associated with the role of the urban environment in the microclimate conditions around the building, building material properties, building equipment usage and HVAC system performance, as well as the occupancy inside the building. These are covered in the UQ repository in the GURA-W.

The proposed baseline model by ESCOMP in the contract adjusts the building's energy consumption for heating (in Therms) based on a regression model that correlates heating demand with weather:

$$\text{Energy}_{\text{heat}} = \text{HDD} \cdot 19.07 - 28{,}250\ ,$$

where HDD is the number of Heating Degree Days in a year. The electricity consumption and demand are not strongly associated with weather and thus need no special treatment. From a mutual agreement between the two parties, any operational and maintenance savings are shared equally, in other words, overall verified savings for this project are as follows:

$$S_V = Cost_B - Cost_P + 0.5 \cdot Cost_M \ ,$$

where $Cost_B - Cost_P$ is the difference between the (adjusted) baseline utility cost and post-retrofit utility bill, and $Cost_M$ is the original average maintenance cost by the school board in Table 7.1. Table 7.1 also shows the corresponding maintenance costs accepted by the ESCOMP, costs associated with measuring and verifying the effective operation of the ECMs, and the installation cost, which includes the cost of any equipment and labor.

Table 7.1. Cost breakdown by energy conservation measure

Energy conservation measure	School board maintenance cost	ESCOMP maintenance cost	Installation cost	M&V cost
Lights and lighting sensors	-$30,900	0[+]	$169,045[+]	$5,000[*]
HVAC set-points and demand ventilation	-$10,500	$6,400	$257,200	$5,000[*]
High efficiency boiler	-$1,450	$450	$57,600	$5,000[*]

[+] Installation cost includes a ten-year maintenance plan that covers all normal maintenance.

[*] If any ECMs are applied, then the total M&V cost is assumed to apply irrespective of ECM choice.

It is assumed that the local government will continue to provide the same level of funding for the duration of the contract. Therefore, if any actual savings were realized by the school board, it would retain those for other purposes. The costs of energy and

operations & maintenance over the past years were on average $179,000. This value is used to estimate the continued operating income received by the school. Finally, we assume ESCOMP assists the school board in arranging a loan from the bank upfront so that the school board has no upfront cost. Therefore, the school board's cash flow during the contract term of ten years is decomposed into annual operating budget, utility bill, remaining maintenance cost if any, mortgage payment to the bank in the amount of guaranteed savings, and potential payout from ESCOMP each year. Likewise, ESCOMP's cash flow comprises upfront payment from the bank less installation cost, and maintenance and M&V cost each year, in addition to potential payout to the school board if guaranteed savings fall short of verified savings. The cost of capital to the school board is modelled as 7%, which translates to a discount rate of 7% in the calculation of net present values. We assume that ESCOMP does not have as easy access to capital, so their discount rate is chosen as 10%.

7.1.2.1 Decision Context

Recall the expected utility theorem, which states that a decision maker should select the alternative with the greatest expected utility. For ESCOMP, as for any ESCO in general, the dominant driver for decision-making is the net present value of profit. However, ESCOMP is not only concerned about the cost and revenue gained from a single retrofit, but is also concerned about the future profitability of the firm. The future profitability is related to the reputation of the firm, as firms that are forced to "pay out" frequently gain the bad reputation of not being able to deliver promised savings. In this study, we assume a current contract could at most influence two other contracts per year in the future, and that the NPV from a contract in the future is the same as the contract under consideration. However, the probability that the ESCOMP receives those future contracts decreases linearly until some critical payout percentage, i.e. 80% in this case, above which no future contracts will be received and ESCOMP risks going out of

business. These factors all come into play when ESCOMP decides on the proposed project scope and the amount of savings to be guaranteed. In terms of the school board, it possesses the ultimate decision authority over whether or not to accept a particular contract. Its fundamental concern is the sustained operation of the school building. Because of the conservative nature of the school board, it is relatively risk averse, whereas ESCOMP is a large company that may decide to accept a larger degree of risk on a given project. Consequently, the school board and ESCOMP's preferences are modeled using a constant absolute risk tolerance of R_C=\$50,000 and R_E=\$100,000, respectively. Another complication is that the client typically charges additional risk premium on top of ESCOMP's estimate of savings uncertainty, resulting from a lack of trust between both parties. We elaborate the above statement as follows. Lee and Paredis (2010) point out that when faced with a future uncertain venture with net present value $V \sim N(\mu, \sigma^2)$, the decision maker seeks to maximize his expected utility:

$$EU = E\left(\mathrm{R} \cdot \left(1 - \mathrm{e}^{-\mathrm{V}/\mathrm{R}}\right)\right) = R \cdot (1 - e^{-\left(\frac{\mu}{R} - \frac{\sigma^2}{2R^2}\right)}) .$$

Note that the $\mathrm{E}(\cdot)$ operator represents expectation with respect to random variable V. It is straightforward to see that maximizing EU is equivalent to maximizing

$$\frac{\mu}{R} - \frac{\sigma^2}{2R^2},$$

as the exponential function e^x monotonically increases with the input x. As such, $\frac{\sigma^2}{2R^2}$ stands for the risk premium term $\mathrm{RP_C}$ and indeed the risk premium increases with an increase in the magnitude of σ^2. In contrast, we assume that a mistrustful client holds the belief that the actual net present value $\mathrm{V}_M = \mu - \sqrt{\alpha} \cdot \sigma|Z|$, where $Z \sim N(0,1)$ and α is the mistrust factor. Such an assumption essentially indicates that a mistrustful client not only increases the magnitude of uncertainty in relation to ESCOMP's estimate, but also assumes zero likelihood of a payback higher than ESCOMP's mean estimate μ. In

mathematical terms, V_M follows a folded normal distribution. As such, for a mistrustful client,

$$EU = E\left(\mathrm{R}\cdot\left(1-\mathrm{e}^{-\mathrm{V}_M/\mathrm{R}}\right)\right) = R\cdot\left(1-e^{-\left(\frac{\mu}{R}-\left(\alpha\cdot\frac{\sigma^2}{2R^2}+\log\left(erfc\left(-\sqrt{\alpha\cdot\frac{\sigma^2}{2R^2}}\right)\right)\right)\right)}\right),$$

where $erfc(\cdot)$ is the Gaussian error function. Similar to the case for a trusting client, maximizing EU for a mistrustful client is equivalent to maximizing

$$\frac{\mu}{R}-\left(\alpha\cdot\frac{\sigma^2}{2R^2}+\log\left(erfc\left(-\sqrt{\alpha\cdot\frac{\sigma^2}{2R^2}}\right)\right)\right).$$

In this case, the risk premium term can be represented by a function of $(\alpha\cdot \mathrm{RP_C})$.

Now we are ready to introduce two scenarios as an illustration of the benefit of transparent risk information. In the first risk-transparent scenario, we assume that ESCOMP and the school board work together, potentially through an independent third party, to analyze the decision context in a transparent manner. With this approach, each party is entitled to their own preferences regarding outcomes, and each describes their beliefs regarding the likelihood of the occurrence of various intermediate outcomes, such that each party can expect the model to yield unbiased and transparent results. Since the school board has the opportunity to more deeply understand and trust the process by which the estimates of energy savings are developed, the board computes its expected utility by remaining loyal to ESCOMP's estimates of NPV distribution. In contrast, the second scenario mimics the business-as-usual case in which no such transparent risk assessment information is available to the client. Expecting a certain level of mistrust by the client, ESCOMP plans for the worst and holds a belief that variable α takes the maximum value of nine. Therefore, in the ESCOMP's worst-case scenario, the most mistrustful client triples the standard deviation of NPV reported by ESCOMP. Figure 7.1 highlights the difference between the two scenarios in terms of assumed NPV distribution.

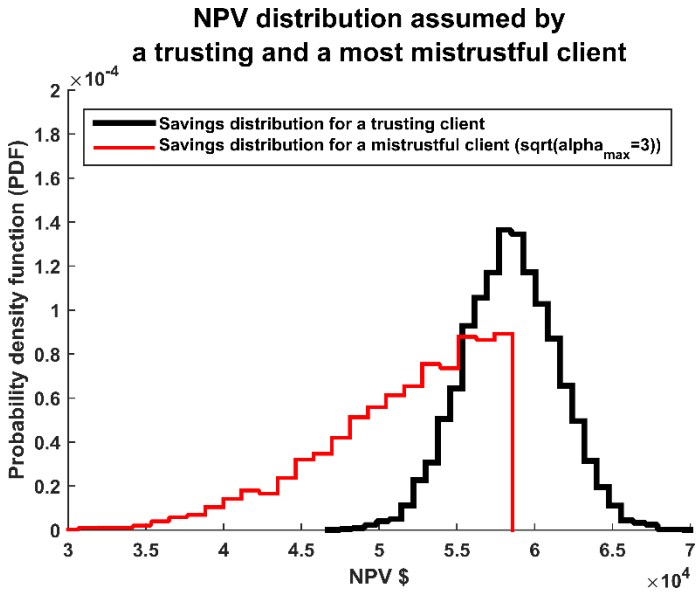


Figure 7.1 Assumed NPV distribution by a trusting and a most mistrustful client

7.1.2.2 Results and Discussions

In a performance contract, ESCOMP first declares a set of ECMs and a guaranteed savings amount, and then the school board determines whether to accept the contract. Given a particular decision by ESCOMP, different potential outcomes occur, and an expectation of utility can be determined for each party. The optimal guaranteed savings to propose by ESCOMP in each alternative is one that maximizes its expected utility, given that the school board is at least better off than its status quo. ESCOMP's final offer to the school board should be the overall best alternative, again measured by expected utility.

The outcomes for both the risk-transparent (RT) case and the business-as-usual (BAU) case are presented in Table 3. There are several common findings between both scenarios. First, it is evident that ESCOMP is willing to propose any alternative in which ECMs are applied, as these alternatives all have a higher expected utility than the status quo in which no ECM is applied. In addition, for some of the alternatives, ESCOMP is willing to take a certain degree of risk, such as a 5% risk of payout, in order to achieve the highest expected utility. Overall, ESCOMP decides to propose the full retrofit

package comprising all three ECMs, as it leads to the highest expected utility. For this package, ESCOMP proposes a guaranteed savings of $84,500, and expects to receive an average NPV of $164,200. The school board approves the proposal as it yields higher expected utility against its status quo. Overall, in most alternatives, implementing the retrofit will help both parties improve over the status quo, which indicates a social surplus and mutual benefit with the acceptance of the contract. Additionally, since the school board is quite risk averse, it weighs potential loss heavier than gain. Hence, from the last column of Table 7.2, we see that the school board is almost indifferent to the choice between the complete package and the "Lights and HVAC" package, because both packages yield significant paybacks. Furthermore, the relationship between guaranteed savings and the expected utility for both ESCOMP and the school board are illustrated in Figure 7.2 for the full retrofit package with a risk-transparent assumption. ESCOMP's expected utility slowly increases with guaranteed savings until it is equivalent to the status quo at a guaranteed savings of approximately $65,500. If guaranteed savings start exceeding this value, eventually payouts begin to occur, and the expected utility increases only slightly at first, then begins to crash suddenly as the loss in reputation heavily affects future revenues. On the other hand, the school board fares best when the guaranteed savings are low. Their profit decreases linearly when the guaranteed savings are met, and then enters a non-linear transitional region, and finally stagnates. In the highest range for guaranteed savings (rightmost part of Figure 7.2), ESCOMP's risk of payout is so high that a large portion of payments is returned to the school board. Such a situation stabilizes the expected utility of the school board.

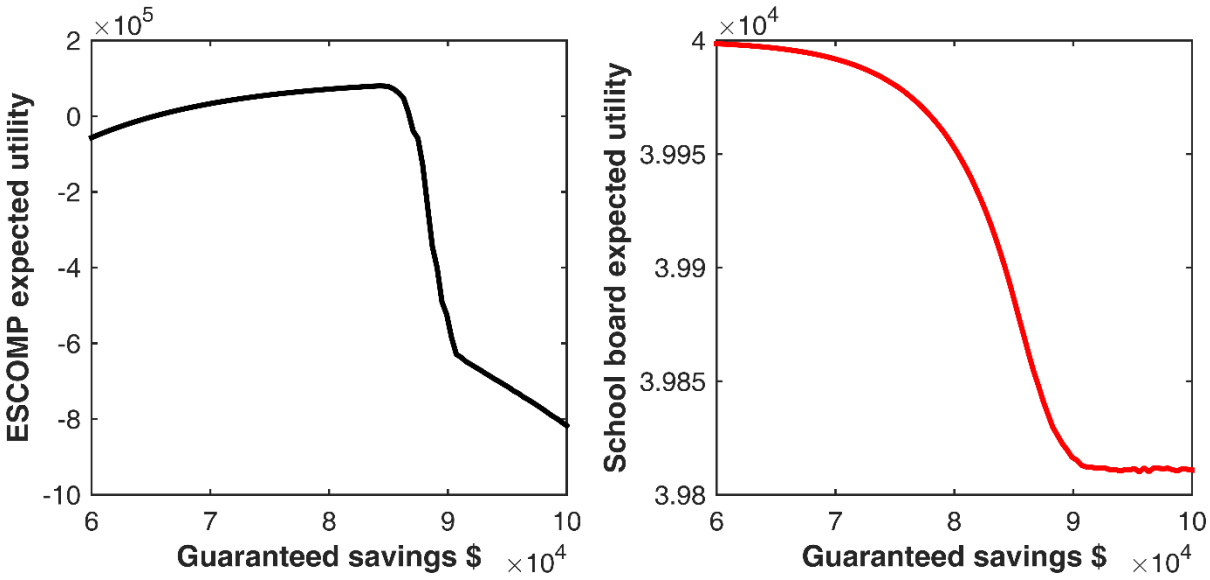


Figure 7.2. Expected utility for ESCOMP (left) and the school board (right)

By contrasting the risk-transparent (RT) and business-as-usual (BAU) case, we also make some interesting discoveries. First, when faced with the same contract terms by ESCOMP, a most mistrustful school board always arrives at a lower expected utility than a trusting one, resulting from a perception of higher risk associated with ESCOMP's estimate of savings. This is a powerful finding of this case study, because it demonstrates that the mistrust due to a lack of risk-transparent information discounts perceived benefits of energy efficiency investments, for the client. In terms of ESCOMP, all optimal guaranteed savings have remained the same except for those of one retrofit package, that is, the "boiler" package. To elaborate, in all other alternatives, although the school board's recognition of benefits have been discounted, its expected utility still improves over the status quo and therefore justifies the acceptance of the contract. This will not affect the optimal guaranteed savings that ESCOMP proposes.

Table 7.2. Results of ECM packages

ECMs	Scenarios	Expected measured savings (k$/yr)	Optimal guaranteed savings (k$/yr)	Expected payout rate (%)	ESCOMP expected NPV profit (k$)	School board expected NPV profit (k$)	ESCOMP expected utility	School board expected utility
Baseline	Common	-	-	-	0.0	58.0	0.0	**30.4**[+]
Lights	RT	36.5	36.1	0.0	108.6	151.3	66.3	39.1
	BAU	36.5	36.1	0.0	108.6	133.5	66.3	38.5
HVAC	RT	45.9	43.6	4.9	39.1	117.3	32.1	37.9
	BAU	45.9	43.8	5.1	39.5	107.2	32.2	37.2
Boiler	RT	17.6	**15.3**	0.0	39.9	65.5	32.9	32.2
	BAU	17.6	**14.8**	0.0	35.4	58.2	29.8	**30.4**
Lights & HVAC	RT	78.9	75.5	0.0	147.7	234.0	77.2	**39.9**
	BAU	78.9	75.5	0.0	147.7	223.3	77.2	39.8
Lights & Boiler	RT	52.3	50.2	1.3	162.3	188.9	80.0	39.6
	BAU	52.3	50.2	1.1	163.3	176.8	80.2	39.5
HVAC & Boiler	RT	53.3	50.6	4.8	38.3	125.9	31.7	38.3
	BAU	53.3	50.6	4.9	38.2	118.6	31.5	37.9
Lights & HVAC & Boiler	RT	87.8	84.5	0	164.2	238.1	**80.6**	**39.9**
	BAU	87.8	84.5	0	164.2	230.3	80.6	**39.9**

[+] The school receives the baseline funding each year based on the pre-retrofit energy consumption records. Since prior years were more severe than predicted future years, the school board would gain a small amount of NPV, even if the status quo persists and no ECMs were applied.

Hence, in most alternatives of this particular case study, the decision outcomes of ESCOMP and the school board are not affected by the mistrust between them. However, when it comes to the "boiler" package, the expect utility of a trusting school board is only barely higher than the status quo. In contrast, a most mistrustful school board further discounts the expected NPV of this package. Therefore, if ESCOMP were to propose the same amount of guaranteed savings at $15,300, the school board would decline the offer. Expecting a mistrustful client in the BAU scenario, ESCOMP has to reduce the amount of guaranteed savings to $14,800 to make the school board indifferent to the acceptance of the project (school board's expected utility breaks even with its status quo), as depicted in Figure 7.3. Consequently, we find that the expected utility of both parties in the BAU scenario have declined against the RT scenario, which essentially indicates a "lose-lose situation" for the "boiler" package. More importantly, in other real-life applications in which an ESCO does not anticipate a high level of mistrust from the client, there is high chance that the project will not be approved eventually. This means a certain level of social surplus will not be materialized without the risk-transparent decision-making framework we propose in this paper.

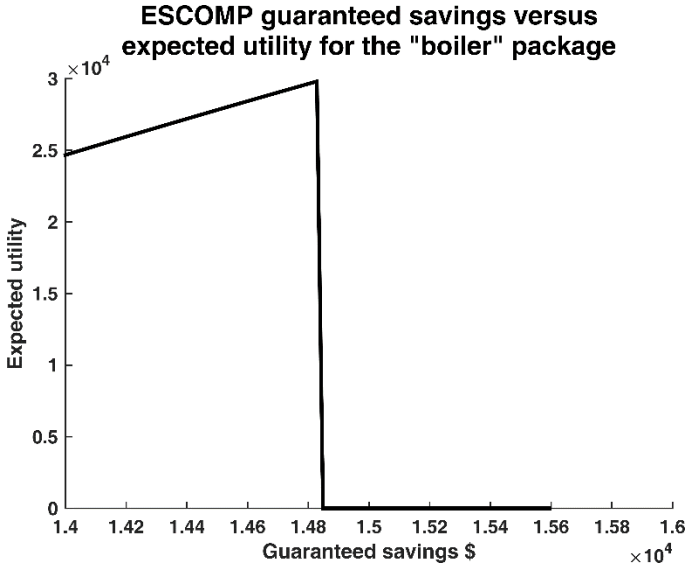


Figure 7.3. ESCOMP's expected utility with proposed guaranteed savings

In summary, our findings are consistent with the more general case described by micro-economic models such as in Allcott and Greenstone (2012). A parameter γ ("implied discount rate") is introduced to represent the tradeoffs decision makers make between upfront investment cost and future savings. In general, the farther parameter γ deviates from 1, the more investment inefficiency such as imperfect information occurs and the heavier the decision maker discounts future energy savings. Our treatment specializes this model by explicitly modeling the increased risk premium charged by the client due to mistrust for the ESCO's estimate of savings uncertainty in an energy performance contract.

CHAPTER 8

EMPIRICAL SPECULATION ABOUT THE RELEVANCE OF UA

This chapter addresses the last question of relevance, and we don't pretend we can give a complete and rigorous answer. The reason for this is that the theoretical basis as well as large-scale data for any type of proof is lacking. However, we do not want to leave the relevance question untouched albeit we will do that briefly in a much less rigorous way. We choose to show an empirical speculation about the relevance of uncertainty analysis, based on a calculation of overall social benefit of a transparent risk quantification framework.

The broader impact of this work can be categorized as an effort for eliminating the "energy efficiency gap" between potential and actually performed energy efficiency projects/savings, created by a variety of market barriers. One of the main hurdles is the lack of established trust in the return on investment of retrofit financing. Despite an abundance of financial incentives and financing models in place, few projects benefit as current protocols cannot handle the large variety of projects each requiring a different risk conscious assessment as the platform for energy saving guarantees and associated performance contracting, which has led to the current situation that the energy efficiency market is too disaggregated to meet the requirement of capital market, such that securitization is not practical or possible. Methodology and methods proposed in this dissertation largely address the above barrier by paving the road towards a protocol for standardized performance risk assessment for energy efficiency projects. Such standardization will streamline the evaluation of energy efficiency investments under uncertainty and promote risk transparency in the negotiation of performance contracts. As stakeholders develop a clear understanding of the associated performance risk, on one hand, a deep retrofit project with low risk will get approved despite its initial rejection

because of the high risk premium charged from the mistrust of projected energy savings; on the other, more informed decisions could be made regarding a project with high risk as well as return because the decision maker could weigh future potential loss versus gain more confidently with the additional risk information.

Conventional research in the energy efficiency space focuses on either increasing the benefit/performance of interventions using novel technology or on decreasing the cost of current technology by improved manufacturing. In contrast, a different focus should be the risk quantification in the energy efficiency market, as advances that pave the road towards reducing perceived risk in energy efficiency investment decisions could potentially encourage as much or greater investments as any particular technology. Therefore, methodology and methods proposed in this dissertation will help scale up and benefit the energy efficiency market as a whole, attracting investments as well as realizing significant energy saving potential that will lead to reductions of imports of energy as well as energy-related emissions.

8.1 Quantification of Impact

Perceived risk with future energy savings stems from the uncertainty associated with the calculation or engineering estimates of savings by current energy simulation tools. Despite their maturity, the prediction of future energy savings remains imprecise, so the modeler has to make assumptions resulting from a lack of information or expertise, which can lead to an overstatement of performance. There is currently no protocol that streamlines a standard approach for the uncertainty quantification and risk assessment of a large variety of projects.

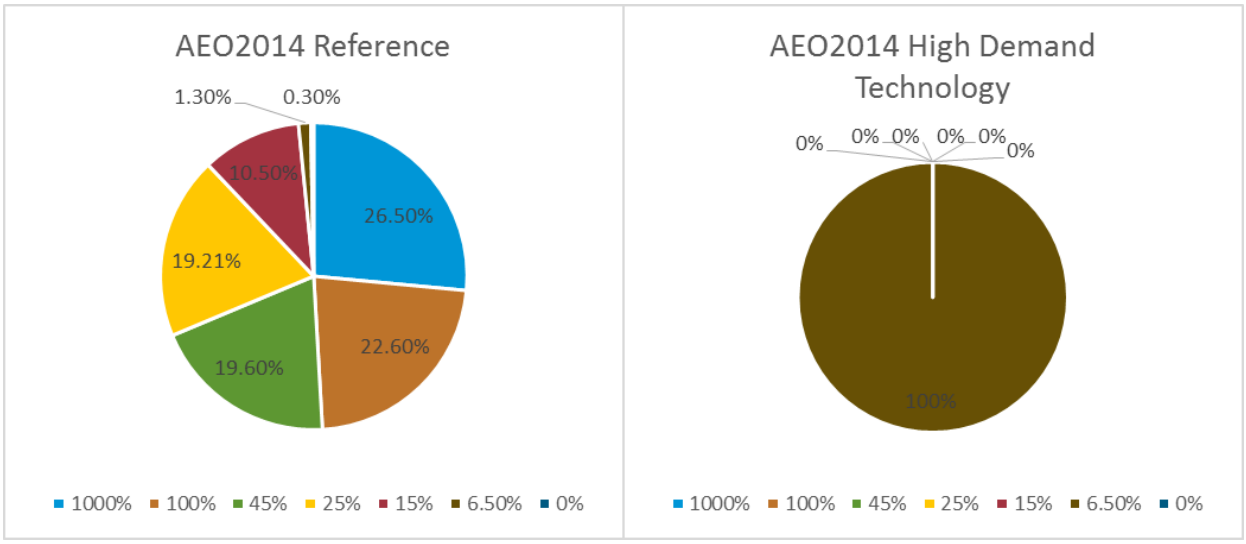


Figure 8.1 Consumer risk premium distribution: baseline (left), high demand (right)

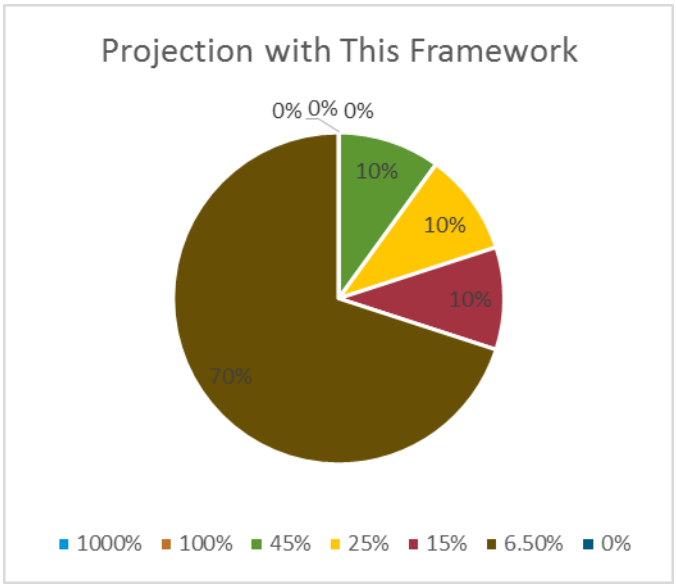


Figure 8.2 Risk premium with risk information: baseline (left), high demand (right)

In order to quantify the broader impact of a transparent risk framework, we seek to the Annual Energy Outlook 2014 published by EIA (Sieminski, 2014) with the underlying National Energy Modeling System (NEMS). The reference/business as usual projection of the commercial delivered energy intensity for 2040 decreases by 16.9% compared to the 2005 baseline. Figure 8.1 shows the distributions of floorspace with the corresponding risk premium from 0% to 1000%. The 1000% premium simulates

floorspace for which existing equipment will never be retrofitted. The business as usual scenario assumes a prescribed distribution for consumer risk-adjusted time preference premium (Figure 8.1 left), which makes up the consumer hurdle (or implicit discount) rates along with the risk free interest rate. Another hypothetical scenario, "High Demand Technology", however, assumes earlier availability, lower costs, and higher efficiencies for more advanced technologies and 6.5% time preference premium for evaluating energy efficiency investments. The transformation from the business as usual scenario to the "High Demand Technology" scenario exactly calls for a standardization approach resulting from methodology and methods proposed in this dissertation, which infuses confidence into the investors and reduces their risk premium. Assuming everything else being equal, we permute the assumption of "High Demand Technology" a bit, depicted in Figure 8.2. Such an assumption reflects the fact that a larger portion of the commercial building stock leaves significant energy saving potential with low risk on the table, while a smaller portion faces the trade-off between future payback and loss. In the former case, the proposed framework enhances investor confidence such that selective NPV-positive EEMs are implemented. In the latter case, buildings identified with high performance risk with certain technology may not be retrofitted despite their approval suggested by conventional methods, which will gradually change the current impression that retrofitted buildings generally under perform their predictions. Identified benefits from the above assumptions are presented in Figure 8.3, which shows that projected delivered energy intensity with transparent risk information is a weighted average between the both EIA AEO 2014 projections, yielding 7.3% savings against the 2040 reference delivered energy intensity. Assuming 1.0% annual growth in commercial floorspace (EIA) with 6.7 Quadrillion BTUs of delivered energy in 2005 as the baseline, the benefits identified above correspond to 0.57 Quadrillion BTUs of energy savings relative to 2040 business as usual scenario. This is a rough estimate of the magnitude of extra savings as a result of uncertainty analysis at large.

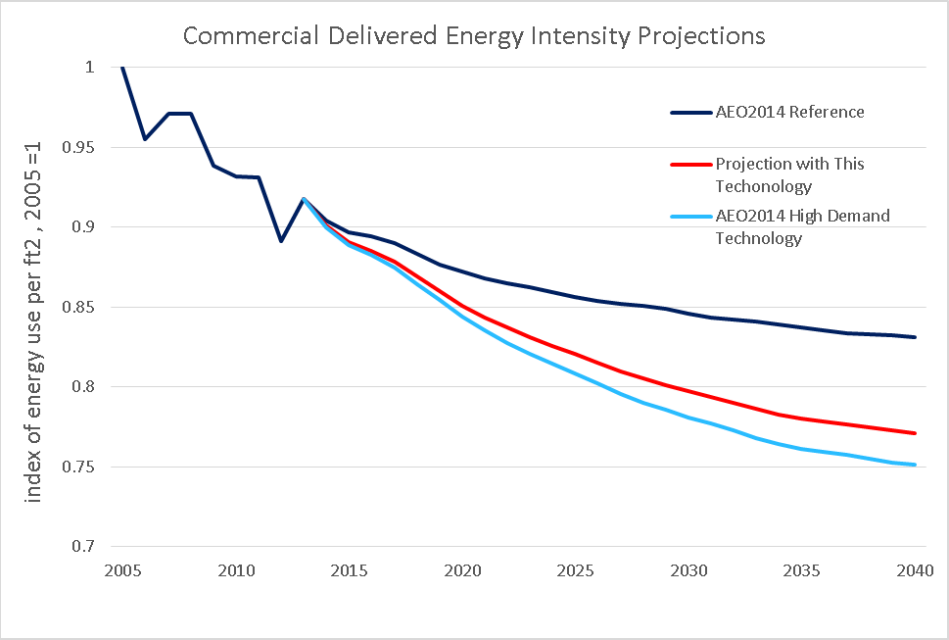


Figure 8.3 Commercial delivered energy intensity projections

Realizing the opportunity identified of course relies on solving several technical challenges, which include an uncertainty analysis workbench for building energy models, a guideline for non-routine adjustment for measurement and verification, a risk conscious decision making framework, and a streamlined process for the management of project financing and implementation, which all combine to form a protocol for energy efficiency standardization. Much of the theoretical foundations have been laid out by this dissertation.

CHAPTER 9

CONCLUSIONS AND FUTURE WORK

9.1 Summary and Conclusions

As new designs push the envelope of building performance, their performance evaluation has to be backed up by building energy simulation. However, various research has shown that deterministic model predictions are anything but assurance that a certain energy performance will be achieved.

This dissertation takes the approach of uncertainty analysis, for a measure of how wrong our predictions are. We build upon previous work and raise important questions of accuracy, validity and relevance for interrogation of our models. The validity question is very important, since without it, we may not really understand how well we are served by these tools, or do they need to be more accurate. Until now, we have addressed each of the three questions in Chapter 1.

Part 1 of the dissertation examined sources of uncertainty not well understood in literature. We started by summarizing common methodology used in uncertainty quantification (UQ) and uncertainty analysis (UA) in Chapter 2, followed by an uncertainty taxonomy and repository in Chapter 3, in preparation for a gap analysis of sources of uncertainty not well studied. In Chapter 4, we built upon previous work and quantified input parameter uncertainty in the well-known infiltration model, model from uncertainty caused by workmanship issues at a thermal bridge, scenario uncertainty by the stochastic usage pattern of occupants, and model from uncertainty from the idealization of the EnergyPlus representation of VAV systems. Furthermore, we proposed improvements over current sensitivity analysis methods, enabling a ranking of model from and scenario uncertainties alongside with input parameter uncertainties.

Part 2 of the dissertation presented a road towards rigorous verification of probabilistic predictions with the aim to prove the ranges of model uncertainties that produce the predictions. In Chapter 6, we advocated for probabilistic and marginal calibration as the corner stone of probabilistic forecast evaluation. The methodology is illustrated by an example that quantifies an aggregated HVAC uncertainty factor that subsumes the effect of model idealization and faults, by careful consideration of other sources of uncertainty.

Part 3 of the dissertation dealt with the validity and relevance proof of this work. We interpreted the validity of uncertainty analysis as leading to "better" decisions. As a preliminary attempt, we defined the best decision as the optimal decision computed by normative utility theory backed up by two key ingredients: carefully elicited risk preferences that reflect decision maker's value, and transparent uncertainty information for energy savings, from an ideal forecast verified by theory proposed in Chapter 6. We then demonstrated how transparent risk information affect decisions in a performance contract for energy retrofits. For the remaining question of relevance, we offered an empirical speculation about the relevance of uncertainty analysis. This was only a speculation as we readily admit that the theoretical basis as well as large-scale data for any hard proof is lacking.

We argue that by addressing the three questions, we add significant knowledge to the building energy research community, with an emphasis on uncertainty analysis.

9.2 Recommendations for Future Study

The ultimate goal of this dissertation, as mentioned in Chapter 8, is to eliminate the energy efficiency gap. This will require that UA and the results of this thesis are adopted in current and future practice. Immediate future work should therefore focus on technology transfer leading to rapid adoption in practice.

First, the expansion of the UQ repository needs to continue, with an emphasis on uncertainty related to HVAC systems. In this dissertation, the top-down approach has only produced a first estimate of the magnitude of HVAC uncertainty, which subsumes HVAC idealization uncertainty and faults which cannot be untangled without further work. Besides, this study is only confined to six buildings on Georgia Tech campus, and we obviously have to repeat similar studies for different pools of buildings. The resulting sets of HVAC uncertainty factors may then be categorized per building age, type, HVAC type etc. As stated previously, the bottom-up approach will require the combined effort of the research community as a whole. This study shows through the example of a VAV system with a particular family of fans and control consequence that the bottom up characterization of HVAC uncertainty can and should be done rigorously. Moreover, the comparison of the top-down macro HVAC factor and the combined uncertainty of bottom up HVAC component uncertainty could prove a very promising way to fully understand the role and origin of HVAC uncertainty and untangle the effects of normal operation uncertainty and incidental faults.

Apart from a focus on HVAC uncertainty, more work should be devoted to identifying those model idealizations that have the highest impact in particular decision-making contexts. This should lead to a prioritization of model improvements in the next generation of building energy software.

Another promising perspective is brought by the advent of the era of big data. We still lack a formal methodology to customize uncertainty analysis for a specific building, with onsite data gathered by the building management systems. The past years have seen an explosion of data collection and storage of data related to building automation, and data analytics is starting to deliver insights that justify the added cost. It may be interesting to find out about the minimum amount of quality data to be collected, for particular combinations of building system and occupancy, in order to de-risk a specific energy efficiency investment. Additionally, we have been interested in the use of

“horizontal” data (consumption data of many buildings but without individual building identity) to make an individual building energy model better. This is an analogy to the classical problem of using general health data to help doctors in diagnosing an individual patient.

The work on performance contracting should be enriched, perhaps with an added flavor from economic theory such as contract and mechanism design. Currently, the mistrust from the clients could originate from the fact that the ESCO can choose not to disclose their private information about how risky a project is for ESCO’s own benefit. Therefore, it will be very interesting to come up with a contract mechanism that advocates a truth-telling behavior for the ESCO. In this dissertation, the decision context is assumed quite simple: ESCO proposes a contract term and the client decides whether to accept the offer, where there is no room for negotiation. Furthermore, besides financial risk, the validity and relevance of uncertainty quantification should also be examined for other scenarios, such as loss of service due to power loss and overheating (Wang, Öcal, Augenbroe, Mengüç, & Özuyar, 2015).

APPENDIX A

PRESSURE LOSS DETAILS FOR THE DUCKWORK STUDIED

Table A.1 Total pressure loss calculations by sections for the actual situation

Duct Section	Duct Element	Airflow, m3/s	Duct Length, m	Pressure Loss, Pa	Section Pressure Loss, Pa
1	Weatherhood			43.3	
	Damper			1.74	
	Filter			134.5	
	Wheel			214	
	Cooling Coil			110.6	
	Heating Coil			17.2	
					521.3
2	Duct	2.313	4.0	2.05	
	Fittings			3.97	
					6.02
3	Duct	1.628	0.5	0.136	
	Fittings			7	
					7.14
4	Duct	1.628	7.3	8.358	
	Fittings			8.71	
					17.07
5	Duct	1.246	7.0	4.854	
	Fittings			82.75	
					87.6
6	Duct	0.316	0.3	0.017	
	Fittings			8.88	

					8.90
7	Duct	0.316	4.3	2.882	
	Fittings			62.04	
					64.92
8	Duct	0.08	1.4	0.082	
	Fittings			13.06	
					13.14
9	Duct	0.08	19.5	8.83	
	Fittings			4.65	
					13.48
10	Duct	0.08	0.6	3.94 *5	
	VAV box			25	
	Diffuser			25	
					69.7
11	Duct	0.236	1.8	1.95	
	Fittings			10.5	
					12.45
12	Duct	0.066	0.9	8.67	
	Fittings			60.8	
					69.47
13	Duct	0.066	0.3	4.29 *5	
	VAV box			25	
	Diffuser			25	
					62.87

Table A.2 Loss coefficient summary by sections for the actual situation

Duct Section	Fitting Number	ASHRAE Fitting No.	Loss Coefficient	Pressure Loss, Pa
2	1	SR7-14	0.08	3.97
3	2	SR5-5	(Cs) 0.1	1.06
	3	SR4-2	0.17	5.94
4	4	CR3-19	0.25	8.71
5	5	SR5-11	(Cs) 0.09	2.42
	6	SR3-1 *2	1.22	24.87 *2
	7	CR9-5	0.46	9.39
	8	SR3-1	1.04	21.2
6	9	SR5-5	(Cs) 4.2	5.52
	10	SR4-1	0.33	3.36
7	11	SR3-1	1.13	11.7
	12	CR3-17	4.43	45.6
	20	CR9-5	0.46	4.74
8	13	SR5-15	(Cb) 13.46	11.66
	14	SR4-3	2.11	1.4
9	15	CD3-13 *3	0.18	0.67 *3
	16	CD3-9 *2	0.23	0.88 *2
	17	CD3-9	0.23	0.88
11	13	SR5-15	(Cs) 0.81	10.5
12	18	SR5-11	(Cb) 1.05	43
	19	CD3-9	0.43	17.8

Table A.3 Total pressure loss calculations by sections for the design situation

Duct Section	Duct Element	Airflow, m3/s	Duct Length, m	Pressure Loss, Pa	Section Pressure Loss, Pa
1	Weatherhood			43.3	
	Damper			1.74	
	Filter			134.5	
	Wheel			214	
	Cooling Coil			110.6	
	Heating Coil			17.2	
					521.3
2	Duct	2.313	4.0	2.05	
	Fittings			3.97	
					6.02
3	Duct	1.628	0.5	0.136	
	Fittings			7	
					7.14
4	Duct	1.628	9.14	10.45	
					10.45
5	Duct	1.246	7.0	4.854	
	Fittings			82.75	
					87.6
6	Duct	0.316	0.3	0.017	
	Fittings			8.88	
					8.90
7	Duct	0.316	4.3	2.882	
	Fittings			42.24	
					45.12

8	Duct	0.08	1.4	0.082	
					0.08
9	Duct	0.08	19.5	8.83	
	Fittings			3.55	
					12.38
10	Duct	0.08	0.6	2.71	
	VAV box			25	
	Diffuser			25	
					52.71
11	Duct	0.066	1.22	0.39	
	Fittings			3.83	
	VAV box			25	
	Diffuser			25	
					54.22
12	Duct	0.165	4.5	4.31	
	Fittings			10.25	
					14.56
13	Duct	0.057	1.22	8.79	
	Fittings			30.85	
	VAV box			25	
	Diffuser			25	
					89.64

Table A.4 Loss coefficient summary by sections for the design situation

Duct Section	Fitting Number	ASHRAE Fitting No.	Loss Coefficient	Pressure Loss, Pa
2	1	SR7-14	0.08	3.97
3	2	SR5-5	(Cs) 0.1	1.06
	3	SR4-2	0.17	5.94
5	6	SR3-1 *2	1.22	24.87 *2
	7	CR9-5	0.46	9.39
	8	SR3-1	1.04	21.2
6	9	SR5-5	(Cs) 4.2	5.52
	10	SR4-1	0.33	3.36
7	11	SR3-1 *3	1.13	11.7 *3
	12	SR5-10	(Cs) 0.83	2.4
	17	CR9-5	0.46	4.74
9	13	SR5-11	(Cs) 0.45	0.39
	14	SR4-3	2.11	1.4
	15	CD3-9 *2	0.23	0.88 *2
11	13	SR5-11	(Cb) 1.49	3.83
12	12	SR5-10	(Cb) 1.62	10.25
13	16	SR5-11	(Cb) 0.81	30.85

REFERENCES

Abushakra, B., & Claridge, D. E. (2008). Modeling Office Building Occupancy in Hourly Data-Driven and Detailed Energy Simulation Programs. *ASHRAE Transactions, 114*(2).

Abushakra, B., Sreshthaputra, A., Haberl, J. S., & Claridge, D. E. (2001). Compilation of Diversity Factors and Schedules for Energy and Cooling Load Calculations, ASHRAE Research Project 1093-RP, Final Report: Energy Systems Laboratory, Texas A&M University.

Abushakra, B., Walker, I. S., & Sherman, M. H. (2004). Compression Effects on Pressure Loss in Flexible HVAC Ducts. *HVAC&R Research, 10*(3), 275-289.

Allcott, H., & Greenstone, M. (2012). Is There an Energy Efficiency Gap? *Journal of Economic Perspectives, 26*(1), 3-28. doi: 10.1257/jep.26.1.3

ASHRAE. (2001). *Handbook of Fundamentals* (Vol. 111): American Society of Heating, Refrigerating and Air Conditioning Engineers, Atlanta.

ASHRAE. (2002). Guideline 14-2002, Measurement of Energy and Demand Savings. Atlanta, Georgia: American Society of Heating, Ventilating, and Air Conditioning Engineers.

ASHRAE. (2005). *Sequences of Operation for Common HVAC Systems*: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

ASHRAE. (2010). Standard 90.1-2010: Energy Standard for Buildings Except Low-Rise Residential Buildings. Atlanta, GA: American Society of Heating, Refrigerating and Air Conditioning Engineers.

ASHRAE. (2011). *Duct Fitting Database*.

Augenbroe, G. (1986). *Research-Oriented Tools for Temperature Calculations in Buildings.* Proc. 2nd Int. Conf. on System Simulation in Buildings, Liege.

Augenbroe, G., Zhang, Y., Khazaii, J., Su, H., Sun, Y., Lee, B. D., & Wu, C. J. (2013). *Implications of the uncoupling of building and HVAC simulation in the presence of parameter uncertainties.* Conference of International Building Performance Simulation Association, Chambery, France.

Aynur, T. N., Hwang, Y. H., & Radermacher, R. (2009a). Simulation Comparison of VAV and VRF Air Conditioning Systems in an Existing Building for the Cooling Season. *Energy and Buildings, 41*(11), 1143-1150. doi: 10.1016/j.enbuild.2009.05.011

Aynur, T. N., Hwang, Y. H., & Radermacher, R. (2009b). Simulation of a VAV Air Conditioning System in an Existing Building for the Cooling Mode. *Energy and Buildings, 41*(9), 922-929. doi: 10.1016/j.enbuild.2009.03.015

Ba, S., Myers, W. R., & Brenneman, W. A. (2015). Optimal Sliced Latin Hypercube Designs. *Technometrics, 57*(4), 479-487. doi: 10.1080/00401706.2014.957867

Bonte, M., Thellier, F., & Lartigue, B. (2014). Impact of Occupant's Actions on Energy Building Performance and Thermal Sensation. *Energy and Buildings, 76*, 219-227. doi: 10.1016/j.enbuild.2014.02.068

Box, G., & Draper, N. (1987). *Empirical model-building and response surfaces*. New York: Wiley.

Bronson, D. J., Hinchey, S. B., Haberl, J. S., & O'Neal, D. L. (1992). *Procedure for Calibrating the DOE-2 Simulation Program to Non-Weather-Dependent*

Measured Loads. Paper presented at the ASHRAE Winter Meeting, Anaheim, CA, USA, 01/25-29/92.
Burhenne, S., Tsvetkova, O., Jacob, D., Henze, G. P., & Wagner, A. (2013). Uncertainty Quantification for Combined Building Performance and Cost-Benefit Analyses. *Building and Environment, 62*, 143-154. doi: 10.1016/j.buildenv.2013.01.013
Chalabi, Z., Das, P., Milner, J., Davies, M., Hamilton, I., Jones, B., . . . Wilkinson, P. (2015). Risk Analysis of Housing Energy Efficiency Interventions Under Model Uncertainty. *Energy and Buildings, 109*, 174-182. doi: 10.1016/j.enbuild.2015.10.006
Chan, W. Y. R., Nazaroff, W. W., Price, P. N., Sohn, M. D., & Gadgil, A. J. (2005). Analyzing a Database of Residential Air Leakage in the United States. *Atmospheric Environment, 39*(19), 3445-3455. doi: 10.1016/j.atmosenv.2005.01.062
Corrado, V., & Mechri, H. E. (2009). Uncertainty and Sensitivity Analysis for Building Energy Rating. *Journal of Building Physics, 33*(2), 125-156. doi: 10.1177/1744259109104884
Daly, D., Cooper, P., & Ma, Z. (2014). Understanding the Risks and Uncertainties Introduced by Common Assumptions in Energy Simulations for Australian Commercial Buildings. *Energy and Buildings, 75*, 382-393. doi: 10.1016/j.enbuild.2014.02.028
de Rocquigny, E., Devictor, N., & Tarantola, S. (2008). *Uncertainty in Industrial Practice: A Guide to Quantitative Uncertainty Management*: John Wiley & Sons.
de Wilde, P. (2014). The Gap Between Predicted and Measured Energy Performance of Buildings: A Framework for Investigation. *Automation in Construction, 41*, 40-49. doi: 10.1016/j.autcon.2014.02.009
de Wilde, P., & Tian, W. (2010). Predicting the Performance of an Office Under Climate Change: A Study of Metrics, Sensitivity and Zonal Resolution. *Energy and Buildings, 42*(10), 1674-1684. doi: 10.1016/j.enbuild.2010.04.011
de Wit, S. (2001). *Uncertainty in Predictions of Thermal Comfort in Buildings*: TU Delft, Delft University of Technology.
de Wit, S., & Augenbroe, G. (2002). Analysis of Uncertainty in Building Design Evaluations and Its Implications. *Energy and Buildings, 34*(9), 951-958. doi: 10.1016/s0378-7788(02)00070-1
Deru, M., & Burns, P. (2003). Infiltration and Natural Ventilation Model for Whole Building Energy Simulation of Residential Buildings. *TRANSACTIONS-AMERICAN SOCIETY OF HEATING REFRIGERATING AND AIR CONDITIONING ENGINEERS, 109*(2), 801-814.
Deru, M., Field, K., Studer, D., Benne, K., Griffith, B., Torcellini, P., . . . Rosenberg, M. (2011). US Department of Energy Commercial Reference Building Models of the National Building Stock: National Renewable Energy Laboratory.
DOE. (2010). Energyplus Engineering Reference.
Dominguez-Munoz, F., Cejudo-Lopez, J. M., & Carrillo-Andres, A. (2010). Uncertainty in Peak Cooling Load Calculations. *Energy and Buildings, 42*(7), 1010-1018. doi: 10.1016/j.enbuild.2010.01.013
Duarte, C., Budwig, R., & Van Den Wymelenberg, K. (2015). Energy and Demand Implication of Using Recommended Practice Occupancy Diversity Factors

Compared to Real Occupancy Data in Whole Building Energy Simulation. *Journal of Building Performance Simulation, 8*(6), 408-423. doi: 10.1080/19401493.2014.966275

Dunn, G., & Knight, I. (2005). Small Power Equipment Loads in UK Office Environments. *Energy and Buildings, 37*(1), 87-91. doi: 10.1016/j.enbuild.2004.05.007

EIA. (2015). How Much Energy Is Consumed in Residential and Commercial Buildings in the United States? , from http://www.eia.gov/tools/faqs/faq.cfm?id=86&t=1

Eisenhower, B., O'Neill, Z., Fonoberov, V. A., & Mezic, I. (2012). Uncertainty and Sensitivity Decomposition of Building Energy Models. *Journal of Building Performance Simulation, 5*(3), 171-184. doi: 10.1080/19401493.2010.549964

Eisenhower, B., O'Neill, Z., Fonoberov, V. A., & Mezić, I. (2012). Uncertainty and Sensitivity Decomposition of Building Energy Models. *Journal of Building Performance Simulation, 5*(3), 171-184. doi: 10.1080/19401493.2010.549964

Emmerich, S. J., & Persily, A. K. (1998). Energy impacts of infiltration and ventilation in US office buildings using multizone airflow simulation. *Proceedings of IAQ and Energy, 98*, 191-206.

Gang, W., Wang, S., Shan, K., & Gao, D. (2015). Impacts of Cooling Load Calculation Uncertainties on the Design Optimization of Building Cooling Systems. *Energy and Buildings, 94*, 1-9. doi: 10.1016/j.enbuild.2015.02.032

Gaterell, M. R., & McEvoy, M. E. (2005). The Impact of Climate Change Uncertainties on the Performance of Energy Efficiency Measures Applied to Dwellings. *Energy and Buildings, 37*(9), 982-995. doi: 10.1016/j.enbuild.2004.12.015

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic Forecasts, Calibration and Sharpness. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 69*, 243-268. doi: 10.1111/j.1467-9868.2007.00587.x

Gneiting, T., & Katzfuss, M. (2014). Probabilistic Forecasting. In S. E. Fienberg (Ed.), *Annual Review of Statistics and Its Application, Vol 1* (Vol. 1, pp. 125-151).

Gneiting, T., & Ranjan, R. (2013). Combining Predictive Distributions. *Electronic Journal of Statistics, 7*, 1747-1782. doi: 10.1214/13-ejs823

Goyal, S., Barooah, P., & Middelkoop, T. (2015). Experimental Study of Occupancy-Based Control of HVAC Zones. *Applied Energy, 140*, 75-84. doi: 10.1016/j.apenergy.2014.11.064

Haves, P., Norford, L. K., DeSimone, M., & Mei, L. (1996). A Standard Simulation Testbed for the Evaluation of Control Algorithms and Strategies, RP-825 Final Report. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Hopfe, C. J., Augenbroe, G. L. M., & Hensen, J. L. M. (2013). Multi-Criteria Decision Making Under Uncertainty in Building Performance Assessment. *Building and Environment, 69*, 81-90. doi: 10.1016/j.buildenv.2013.07.019

Hopfe, C. J., & Hensen, J. L. M. (2011). Uncertainty Analysis in Building Performance Simulation for Design Support. *Energy and Buildings, 43*(10), 2798-2805. doi: 10.1016/j.enbuild.2011.06.034

Howard, R. A. (1988). Decision-Analysis - Practice and Promise. *Management Science, 34*(6), 679-695. doi: 10.1287/mnsc.34.6.679

Huang, P., Huang, G., & Wang, Y. (2015). HVAC system design under peak load prediction uncertainty using multiple-criterion decision making technique. *Energy and Buildings, 91*, 26-36. doi: 10.1016/j.enbuild.2015.01.026

Huang, P., Wang, Y., Huang, G., & Augenbroe, G. (2015). Investigation of the Ageing Effect on Chiller Plant Maximum Cooling Capacity Using Bayesian Markov Chain Monte Carlo Method. *Journal of Building Performance Simulation*, 1-13. doi: 10.1080/19401493.2015.1117529

IEA-EBC Annex 66. (2015). *Introduction.* Retrieved from http://www.annex66.org/

ISO. (2007). 14683: Thermal Bridges in Building Construction–Linear Thermal Transmittance–Simplified Methods and Default Values (ISO 14683: 2007): CEN.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 112): Springer.

Jokisalo, J., Kurnitski, J., Korpi, M., Kalamees, T., & Vinha, J. (2009). Building Leakage, Infiltration, and Energy Performance Analyses for Finnish Detached Houses. *Building and Environment, 44*(2), 377-387. doi: 10.1016/j.buildenv.2008.03.014

Judkoff, R., Wortman, D., O'Doherty, B., & Burch, J. (2008). A Methodology for Validating Building Energy Analysis Simulations: National Renewable Energy Laboratory.

Kalamees, T., Korpi, M., Eskola, L., Kurnitski, J., & Vinha, J. (2008). *The Distribution of the Air Leakage Places and Thermal Bridges in Finnish Detached Houses and Apartment Buildings.* Proceedings of the 8th Symposium on Building Physics in the Nordic Countries (NSB2008), Copenhagen, Denmark.

Kerwin, A. (1993). None Too Solid - Medical Ignorance. *Knowledge-Creation Diffusion Utilization, 15*(2), 166-185.

Kim, S. H., & Augenbroe, G. (2013). Uncertainty in Developing Supervisory Demand-Side Controls in Buildings: A Framework and Guidance. *Automation in Construction, 35*, 28-43. doi: 10.1016/j.autcon.2013.02.001

Lazzarin, R. M. (2012). Condensing Boilers in Buildings and Plants Refurbishment. *Energy and Buildings, 47*, 61-67. doi: 10.1016/j.enbuild.2011.11.029

LBNL Simulation Research Group. (2016). FMU Export of EnergyPlus.

Lee, B. D. (2014). *A Pragmatic Value-Driven Approach to Design with Applications to Energy-Conscious Buildings (Doctoral Dissertation).* Retrieved from http://hdl.handle.net/1853/53094 Georgia Tech Theses and Dissertations database.

Lee, B. D., & Paredis, C. J. J. (2010). Accounting for the Duration of Analyses in Design Process Decisions. *SAE International Journal of Materials & Manufacturing*, 3(1):512-522. doi: 10.4271/2010-01-0908

Li, Z., Heo, Y., & Augenbroe, G. (2009). *HVAC Design Informed by Organizational Simulation.* Building Simulation.

Liao, Y., Huang, G., Sun, Y., & Zhang, L. (2014). Uncertainty Analysis for Chiller Sequencing Control. *Energy and Buildings, 85*, 187-198. doi: 10.1016/j.enbuild.2014.09.037

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS Book: A Practical Introduction to Bayesian Analysis*: CRC press.

Maasoumy, M., Razmara, M., Shahbakhti, M., & Vincentelli, A. S. (2014). Handling Model Uncertainty in Model Predictive Control for Energy Efficient Buildings. *Energy and Buildings, 77*, 377-392. doi: 10.1016/j.enbuild.2014.03.057

Macdonald, I. (2002). *Quantifying the Effects of Uncertainty in Building Simulation*: University of Strathclyde.

Macdonald, I., & Strachan, P. (2001). Practical Application of Uncertainty Analysis. *Energy and Buildings, 33*(3), 219-227. doi: 10.1016/s0378-7788(00)00085-2

Mahdavi, A. (2010). People in Building Performance Simulation. In J. Hensen & R. Lamberts (Eds.), *Building Performance Simulation for Design and Operation* (pp. 56-83).

Mara, T. A., & Tarantola, S. (2008). Application of Global Sensitivity Analysis of Model Output to Building Thermal Simulations. *Building Simulation, 1*(4), 290-302. doi: 10.1007/s12273-008-8129-5

Martin, K., Escudero, C., Erkoreka, A., Flores, I., & Sala, J. M. (2012). Equivalent Wall Method for Dynamic Characterisation of Thermal Bridges. *Energy and Buildings, 55*, 704-714. doi: 10.1016/j.enbuild.2012.08.024

Matheson, J. E., & Winkler, R. L. (1976). Scoring Rules for Continuous Probability Distributions. *Management Science, 22*(10), 1087-1096. doi: 10.1287/mnsc.22.10.1087

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A Comparision of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics, 21*(2), 239-245. doi: 10.2307/1268522

Mechri, H. E., Capozzoli, A., & Corrado, V. (2010). USE of the ANOVA Approach for Sensitive Building Energy Design. *Applied Energy, 87*(10), 3073-3083. doi: 10.1016/j.apenergy.2010.04.001

Mei, L., & Levermore, G. J. (2002). Simulation and Validation of a VAV System with an ANN Fan Model and a Non-Linear VAV Box Model. *Building and Environment, 37*(3), 277-284. doi: 10.1016/s0360-1323(01)00028-2

Meier, L. (2015). grplasso: Fitting user specified models with Group Lasso penalty. *R package version 0.4-5*. from https://CRAN.R-project.org/package=grplasso

Menassa, C. C. (2011). Evaluating Sustainable Retrofits in Existing Buildings Under Uncertainty. *Energy and Buildings, 43*(12), 3576-3583. doi: 10.1016/j.enbuild.2011.09.030

Molina, M. (2014). The Best Value for America's Energy Dollar: A National Review of the Cost of Utility Energy Efficiency Programs *American Council for an Energy-Efficient Economy*.

Moon, H. J. (2005). *Assessing Mold Risks in Buildings Under Uncertainty (Doctoral Dissertation).* Retrieved from http://hdl.handle.net/1853/7279 Georgia Tech Theses and Dissertations database.

Morgan, M. G. (2009). *Best Practice Approaches for Characterizing, Communicating and Incorporating Scientific Uncertainty in Climate Decision Making*: DIANE Publishing.

National Lighting Product Information Program (NLPIP). (1999). *National Lighting Product Information Program: Specifier Reports.* Retrieved from http://www.lrc.rpi.edu/programs/nlpip/screwbase.asp

Neumann, V. J., & Morgenstern, O. (1944). Theory of Games and Economic Behavior. *Princeton University Prress, Princeton*.

Nouidui, T. S. (2014, August 26-28, 2013). *Functional Mock-Up Unit Import in Energyplus for Co-Simulation.* 13th Conference of International Building Performance Simulation Association, Chambéry, France.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., . . . Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*: John Wiley & Sons.

Oberkampf, W. L., & Roy, C. J. (2010). *Verification and Validation in Scientific Computing*: Cambridge University Press.

Oldewurtel, F., Sturzenegger, D., & Morani, M. (2013). Importance of Occupancy Information for Building Climate Control. *Applied Energy, 101*, 521-532. doi: 10.1016/j.apenergy.2012.06.014

Pappenberger, F., Iorgulescu, I., & Beven, K. J. (2006). Sensitivity Analysis Based on Regional Splits and Regression Trees (SARS-RT). *Environmental Modelling & Software, 21*(7), 976-990. doi: 10.1016/j.envsoft.2005.04.010

Parys, W., Breesch, H., Hens, H., & Saelens, D. (2012). Feasibility Assessment of Passive Cooling for Office Buildings in a Temperate Climate Through Uncertainty Analysis. *Building and Environment, 56*, 95-107. doi: 10.1016/j.buildenv.2012.02.018

Parys, W., Saelens, D., & Hens, H. (2011). Coupling of Dynamic Building Simulation with Stochastic Modelling of Occupant Behaviour in Offices - a Review-Based Integrated Methodology. *Journal of Building Performance Simulation, 4*(4), 339-358. doi: 10.1080/19401493.2010.524711

Pazold, M., Burhenne, S., Radon, J., Herkel, S., & Antretter, F. (2012, September 3-5, 2012). *Integration of Modelica Models into an Existing Simulation Software Using FMI for Co-Simulation.* 9th International MODELICA Conference, Munich, Germany.

Perez, R., Stewart, R., Arbogast, C., Seals, R., & Scott, J. (1986). An Anisotropic Hourly Diffuse-Radiation Model for Sloping Surfaces-Description, Performance Validation, Site Dependency Evaluation. *Solar Energy, 36*(6), 481-497. doi: 10.1016/0038-092x(86)90013-7

Persily, A. K. (1998). *Airtightness of Commercial and Institutional Buildings: Blowing Holes in the Myth of Tight Buildings*.

PHYSIBEL. (2002). *KOBRA: User's guide*. PHYSIBEL.

Purdy, J., & Beausoleil-Morrison, I. (2001). *The Significant Factors in Modelling Residential Buildings.* Seventh International IBPSA Conference, RIo de Janeiro, Brazil, August 13.

Rasouli, M., Ge, G. M., Simonson, C. J., & Besant, R. W. (2013). Uncertainties in Energy and Economic Performance of HVAC Systems and Energy Recovery Ventilators Due to Uncertainties in Building and HVAC Parameters. *Applied Thermal Engineering, 50*(1), 732-742. doi: 10.1016/j.applthermaleng.2012.08.021

Reinhart, C. F. (2004). Lightswitch-2002: a Model for Manual and Automated Control of Electric Lighting and Blinds. *Solar Energy, 77*(1), 15-28. doi: 10.1016/j.solener.2004.04.003

Reinhold, C., & Sonderegger, R. (1983). Component Leakage Areas in Residential Buildings: Lawrence Berkeley Lab., CA (USA).
Roache, P. J. (2009). *Fundamentals of Verification and Validation*: Hermosa publishers Socorro.
Roberson, J. A., Webber, C. A., McWhinney, M. C., Brown, R. E., Pinckard, M. J., & Busch, J. F. (2004). After-Hours Power Status of Office Equipment and Inventory of Miscellaneous Plug-Load Equipment. *Lawrence Berkeley National Laboratory*.
Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics, 23*(3), 470-472.
Ruiz, F. R., Bertagnolio, S., & Lemort, V. (2012). *Global Sensitivity Analysis applied to Total Energy Use in Buildings.* Proceedings of the 2nd International High Performance Buildings Conference.
Ryan, E. M., & Sanquist, T. F. (2012). Validation of Building Energy Modeling Tools Under Idealized and Realistic Conditions. *Energy and Buildings, 47*, 375-382. doi: 10.1016/j.enbuild.2011.12.020
Rysanek, A. M., & Choudhary, R. (2015). DELORES - an Open-Source Tool for Stochastic Prediction of Occupant Services Demand. *Journal of Building Performance Simulation, 8*(2), 97-118. doi: 10.1080/19401493.2014.888595
Saelens, D., Parys, W., & Baetens, R. (2011). Energy and Comfort Performance of Thermally Activated Building Systems Including Occupant Behavior. *Building and Environment, 46*(4), 835-848. doi: 10.1016/j.buildenv.2010.10.012
Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*: John Wiley & Sons.
Schweigler, L. M., Desmond, J. S., McCarthy, M. L., Bukowski, K. J., Ionides, E. L., & Younger, J. G. (2009). Forecasting Models of Emergency Department Crowding. *Academic Emergency Medicine, 16*(4), 301-308. doi: 10.1111/j.1553-2712.2009.00356.x
Sherman, M. H., & Grimsrud, D. T. (1980). Measurement of Infiltration Using Fan Pressurization and Weather Data: Lawrence Berkeley Lab., CA (USA).
Sieminski, A. (2014). International Energy Outlook: Energy Information Administration (EIA).
Silva, A. S., & Ghisi, E. (2014a). Uncertainty Analysis of the Computer Model in Building Performance Simulation. *Energy and Buildings, 76*, 258-269. doi: 10.1016/j.enbuild.2014.02.070
Silva, A. S., & Ghisi, E. (2014b). Uncertainty Analysis of User Behaviour and Physical Parameters in Residential Building Performance Simulation. *Energy and Buildings, 76*, 381-391. doi: 10.1016/j.enbuild.2014.03.001
Smith, A., Luck, R., & Mago, P. J. (2010). Analysis of a Combined Cooling, Heating, and Power System Model Under Different Operating Strategies with Input and Model Data Uncertainty. *Energy and Buildings, 42*(11), 2231-2240. doi: 10.1016/j.enbuild.2010.07.019
Spitz, C., Mora, L., Wurtz, E., & Jay, A. (2012). Practical Application of Uncertainty Analysis and Sensitivity Analysis on an Experimental House. *Energy and Buildings, 55*, 459-470. doi: 10.1016/j.enbuild.2012.08.013

Srivastav, A., Tewari, A., & Dong, B. (2013). Baseline Building Energy Modeling and Localized Uncertainty Quantification Using Gaussian Mixture Models. *Energy and Buildings, 65*, 438-447. doi: 10.1016/j.enbuild.2013.05.037

Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics, 29*(2), 143-151. doi: 10.2307/1269769

Sun, Y. M. (2014). *Closing the Building Energy Performance Gap by Improving Our Predictions (Doctoral Dissertation).* Retrieved from http://hdl.handle.net/1853/52285 Georgia Tech Theses and Dissertations database.

Sun, Y. M., Gu, L., Wu, C. F. J., & Augenbroe, G. (2014). Exploring HVAC System Sizing Under Uncertainty. *Energy and Buildings, 81*, 243-252. doi: 10.1016/j.enbuild.2014.06.026

Sun, Y. M., Heo, Y., Tan, M. H. Y., Xie, H. Z., Wu, C. F. J., & Augenbroe, G. (2014). Uncertainty Quantification of Microclimate Variables in Building Energy Models. *Journal of Building Performance Simulation, 7*(1), 17-32. doi: 10.1080/19401493.2012.757368

Sun, Y. M., Su, H., Wu, C. F. J., & Augenbroe, G. (2015). Quantification of Model Form Uncertainty in the Calculation of Solar Diffuse Irradiation on Inclined Surfaces for Building Energy Simulation. *Journal of Building Performance Simulation, 8*(4), 253-265. doi: 10.1080/19401493.2014.914247

Swan, L., Ugursal, V. I., & Beausoleil-Morrison, I. (2009). Implementation of a Canadian Residential Energy End-Use Model for Assessing New Technology Impacts. *Proceedings of Building Simulation, Glasgow*, 1429-1436.

Tahmasebi, F., & Mahdavi, A. (2015). The Sensitivity of Building Performance Simulation Results to the Choice of Occupants' Presence Models: A Case Study. *Journal of Building Performance Simulation*, 1-11. doi: 10.1080/19401493.2015.1117528

Tian, W. (2013). A Review of Sensitivity Analysis Methods in Building Energy Analysis. *Renewable and Sustainable Energy Reviews, 20*, 411-419. doi: http://dx.doi.org/10.1016/j.rser.2012.12.014

Tian, W., & de Wilde, P. (2011). Uncertainty and Sensitivity Analysis of Building Performance Using Probabilistic Climate Projections: A UK Case Study. *Automation in Construction, 20*(8), 1096-1109. doi: 10.1016/j.autcon.2011.04.011

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B-Methodological, 58*(1), 267-288.

Toth, Z., Talagrand, O., Candille, G., & Zhu, Y. (2003). Forecast Verification: A Practitioner's Guide in Atmospheric Science: Wiley, Chapter Probability and ensemble forecasts.

Turner, C., & Frankel, M. (2008). Energy Performance of LEED for New Construction Buildings. *New Buildings Institute, 4*, 1-42.

Turner, C., Frankel, M., & Council, U. S. G. B. (2008). *Energy Performance of LEED for New Construction Buildings*: New Buildings Institute Vancouver, WA.

U.S. Green Building Council. (2015). Green Building Facts.

Walker, I. S., & Wilson, D. J. (1993). Evaluating Models for Superposition of Wind and Stack Effect in Air Infiltration. *Building and Environment, 28*(2), 201-210. doi: 10.1016/0360-1323(93)90053-6

Walker, I. S., & Wilson, D. J. (1998). Field Validation of Algebraic Equations for Stack and Wind Driven Air Infiltration Calculations. *HVAC&R Research, 4*(2), 119-139.

Wang, L. (2014). *Modeling and Simulation of HVAC Faulty Operations and Performance Degradation Due to Maintenance Issues.* ASim 2012-1st Asia conference of International Building Performance Simulation Association, Shanghai, China, 11/25/12-11/27/12.

Wang, L. P., Mathew, P., & Pang, X. F. (2012). Uncertainties in Energy Consumption Introduced by Building Operations and Weather for a Medium-Size Office Building. *Energy and Buildings, 53*, 152-158. doi: 10.1016/j.enbuild.2012.06.017

Wang, Q., Augenbroe, G., Kim, J., & Gu, L. (2016). Meta-Modeling of Occupancy Variables and Analysis of Their Impact on Energy Outcomes of Office Buildings. *Applied Energy, 174*, 166-180. doi: http://dx.doi.org/10.1016/j.apenergy.2016.04.062

Wang, Q., Augenbroe, G., & Sun, Y. (2014). *The Role of Construction Detailing and Workmanship in Achieving Energy Efficient Buildings.* Construction Research Congress, Atlanta, May.

Wang, Q., Öcal, M. R., Augenbroe, G., Mengüç, M. P., & Özuyar, P. (2015). *An Evaluation of Energy Efficiency Measures in a Turkish Campus Building for Thermal Comfort and Economic Risk*. Paper presented at the Building Simulation Conference 2015, Hyderabad, India.

Wang, S. W., & Jin, X. Q. (2000). Model-Based Optimal Control of VAV Air-Conditioning System Using Genetic Algorithm. *Building and Environment, 35*(6), 471-487. doi: 10.1016/s0360-1323(99)00032-3

Wang, W. M., Beausoleil-Morrison, I., & Reardon, J. (2009). Evaluation of the Alberta Air Infiltration Model Using Measurements and Inter-Model Comparisons. *Building and Environment, 44*(2), 309-318. doi: 10.1016/j.buildenv.2008.03.005

Wetter, M., Zuo, W. D., Nouidui, T. S., & Pang, X. F. (2014). Modelica Buildings Library. *Journal of Building Performance Simulation, 7*(4), 253-270. doi: 10.1080/19401493.2013.765506

Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences* (Vol. 100): Academic press.

Wray, C. P., Diamond, R. C., & Sherman, M. H. (2005). Rationale for Measuring Duct Leakage Flows in Large Commercial Buildings. *Lawrence Berkeley National Laboratory*.

Wray, C. P., & Matson, N. E. (2003). Duct Leakage Impacts on VAV System Performance in California Large Commercial Buildings. *Lawrence Berkeley National Laboratory*.

Yan, B. (2016). Quantifying Uncertainty in Outdoor Air Flow Control and its Impacts on Building Performance Simulation and Fault Detection. *(to be published)*.

Yang, X. B., Jin, X. Q., Du, Z. M., Fan, B., & Chai, X. F. (2011). Evaluation of Four Control Strategies for Building VAV Air-Conditioning Systems. *Energy and Buildings, 43*(2-3), 414-422. doi: 10.1016/j.enbuild.2010.10.004

Yoshida, H., Kumar, S., & Morita, Y. (2001). Online Fault Detection and Diagnosis in VAV Air Handling Unit by RARX Modeling. *Energy and Buildings, 33*(4), 391-401. doi: 10.1016/s0378-7788(00)00121-3

Yuan, M., & Lin, Y. (2006). Model Selection and Estimation in Regression with Grouped Variables. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 68*, 49-67. doi: 10.1111/j.1467-9868.2005.00532.x